

**Which?**

Accent

 PJM economics

RESEARCH REPORT SEPTEMBER 2021

---

# Value of the Choice Requirement Remedy

# Contents

<b>About</b>	<b>4</b>
<b>Executive Summary</b>	<b>5</b>
<b>1 Introduction</b>	<b>8</b>
1.1 Background	8
1.2 Objectives	8
1.3 Report Structure	8
<b>2 Context for the Study</b>	<b>10</b>
2.1 Policy Context	10
2.2 Privacy Concerns	13
2.3 Valuations of Privacy	15
<b>3 Methodology</b>	<b>19</b>
3.1 Introduction	19
3.2 Contingent Valuation Design	19
3.3 Questionnaire Structure	25
3.4 Development and Testing	25
3.5 Main Survey Administration	26
3.6 Data Diagnostics and Sensitivity Samples	27
<b>4 Usage, Awareness and Attitudes</b>	<b>29</b>
4.1 Introduction	29
4.2 Service and Device Usage	29
4.3 Awareness of and Attitudes towards Targeted Adverts	34
4.4 Privacy Settings	37
4.5 Informed Attitudes towards Data Sharing	40

<b>5</b>	<b>Contingent Valuation Analysis and Findings</b>	<b>43</b>
5.1	Introduction	43
5.2	Descriptive Analysis	43
5.3	Valuation of the Choice Requirement Remedy	52
5.4	Choices Between Data Sharing Alternatives	57
<b>6</b>	<b>Conclusions</b>	<b>61</b>
	<b>References</b>	<b>63</b>
	<b>APPENDIX A Main Survey Questionnaire</b>	<b>69</b>
	<b>APPENDIX B Sample Demographics</b>	<b>88</b>
	<b>APPENDIX C Econometric Modelling of Contingent Valuation Choices</b>	<b>90</b>
	<b>APPENDIX D Literature Review Estimates</b>	<b>106</b>
	<b>APPENDIX E Review by Prof. Ken Willis</b>	<b>108</b>

# About

## Which?

Which? is the UK's consumer champion. As an organisation we're not for profit - a powerful force for good, here to make life simpler, fairer and safer for everyone. We're the independent consumer voice that provides impartial advice, investigates, holds businesses to account and works with policymakers to make change happen. We fund our work mainly through member subscriptions, we're not influenced by third parties and we buy all the products that we test.

## Accent

Accent is a full service quantitative and qualitative market research agency with over 30 years of experience. We provide our clients with bespoke research solutions for strategic issues. We apply fresh and innovative thinking, and approach every brief with a fresh perspective using ground breaking and award winning quantitative and qualitative approaches.

## PJM Economics

PJM Economics is a small team of microeconomists led by founding director Dr Paul Metcalfe. The company provides bespoke economic research and consultancy services primarily in sectors subject to economic regulation. Core service areas include, economic valuation, appraisal and the modelling and forecasting of microeconomic variables

# Executive Summary

## Introduction

In July 2020, the Competition and Markets Authority (CMA) published a market study focused on Online Platforms and Digital Advertising, which concluded that consumers were receiving inadequate compensation for the use of their personal data by online platforms. On the basis of these conclusions, the CMA recommended that the regulator be given powers to introduce the following two key interventions:

- *The choice requirement remedy*: requiring platforms to give consumers the choice not to share their data for personalised advertising; and,
- *The 'Fairness by Design' duty*: placing a duty on platforms to ensure that they are maximising users' ability to make informed choices about the use of their personal data.

The objective for the present research was to obtain user valuations for the CMA's proposed 'Choice requirement remedy' (CRR) via a stated preference methodology. This value evidence was intended to contribute to the broader evidence base in this area, with a view that it could potentially inform government decisions around potential legislation.

## Survey Design and Administration

A stated preference survey methodology was implemented, involving a large UK survey of Facebook / Google users. The survey instrument was designed to value the CRR based on the premise that it effectively adds a new option for users of Facebook and Google regarding the sharing of their data for the purposes of targeted advertising. Our approach accordingly assumes a zero value for the CRR for all users that had already successfully reviewed their settings, whether or not they made any changes, on the basis that it does not add a new option for these people. This is a cautious approach, which tends to understate the value of the CRR, as users who have already reviewed their settings may experience a welfare gain from a more accessible and wider choice between data sharing options following implementation of the CRR.

The survey included a sequence of contingent valuation questions asking participants to choose between receiving 'Generic' adverts (no data sharing) and 'Targeted' adverts (full data sharing). In addition, one further question offered a choice of multiple alternatives for the type of personal data that would be shared.

Participants were randomly assigned to one of two 'treatments', which carried through across all of these valuation questions:

- A fee to be paid for receiving 'Generic' adverts only; or,
- A reward to be earned for receiving 'Targeted' adverts.

The reward scenario was believed to be more likely following implementation of the CRR, as envisaged by the CMA, but it represents a less conservative approach to valuation in comparison with a fee scenario. Our main estimates are accordingly based on the conservative ‘fee’ scenario.

The contingent valuation questions were asked both prior to, and following, the presentation of detailed information regarding the type of information that was used by Facebook/Google for the purposes of showing targeted adverts. This approach results in both ‘uninformed’ and ‘informed’ valuations, and thereby enables the impact of information to be examined. While uninformed choices provide a point of comparison for assessing the impact of information on the valuation of data privacy, the informed choices are those relevant for the valuation of the CRR, given that the complementary ‘Fairness by design’ intervention is aimed at ‘maximising users’ ability to make informed choices about the use of their personal data’ (CMA, 2020a). Hence, our main estimates are based on the informed scenario.

The main survey was conducted online from 19 May to 1 June 2021, following several phases of development and testing research. The sample was sourced from an online panel, and a total of 4,014 fully completed interviews were achieved.

It is possible that online panel members might have a greater willingness to share data about themselves than the general population of users, on account of them having signed up to complete surveys on a regular basis. If this is the case, the valuations obtained may prove to be a conservative underestimate of the true population valuations.

## Key findings

The main valuation findings are shown in Table 1. Facebook and Google results are combined due to the fact there were no statistically significant differences between services. The table shows that mean valuations of the CRR were higher for informed choices than uninformed choices, indicating that the information shown caused people to express a stronger preference for the privacy-enhanced option. Valuations were also higher under the reward scenario than under the fee scenario, as was expected given previous findings on similar comparisons.

Under the preferred fee/informed scenario, the estimated value is £1.09 per user per month. This implies a total valuation of the CRR in the UK, as applied to Facebook and Google, of £1,136m per year.

**Table 1:** Valuations of the CMA’s proposed Choice Requirement Remedy (Applied to Facebook and Google)

	Fee scenario		Reward scenario	
	Uninformed choices	Informed choices	Uninformed choices	Informed choices
Mean valuation per Facebook/Google user (£/month)	£0.50 [£0.41–£0.59]	<b>£1.09</b> <b>[£0.95–£1.23]</b>	£1.68 [£1.38–£1.99]	£4.03 [£3.26–£4.79]
Total valuation by UK internet users (£m/year)	£527m [£434m–£620m]	<b>£1,136m</b> <b>[£989m–£1,283m]</b>	£1,761m [£1,441m–£2,082m]	£4,209m [£3,409m–£5,009m]

95% confidence intervals in parentheses. See Table 12, Table 13, and surrounding text for methodological details.

There are several indications supporting the validity and reliability of the valuations:

- An extensive development and testing programme found that the core choices were considered plausible by participants, and that participants had no problems answering these questions appropriately.
- Values were obtained from a large national sample of Google and Facebook users.
- A conservative approach was taken to valuation, including adopting the lower fee-based values rather than the higher reward-based estimates. Additionally, the use of an online panel, rather than a face-to-face or telephone sample, can also be expected to have produced conservative valuations.
- Valuations varied in line with expectation. For example, those saying they were less comfortable sharing their data were found to accordingly have higher valuations than those more comfortable sharing their data.
- Values are consistent with other results reported in the literature on privacy valuation.

There are hence good grounds for considering the valuations of the CRR to be robust and reliable for users of Facebook and Google in the UK.

# 1. Introduction

## 1.1 Background

In July 2020, the Competition and Markets Authority (CMA) published a market study focused on Online Platforms and Digital Advertising<sup>1</sup>. This study assessed the markets for search, social media and digital advertising, and the role of Google and Facebook within them. It concluded that Google and Facebook enjoy substantial market power in the search and display advertising markets respectively, and that this results in consumers receiving inadequate compensation for their attention and the use of their personal data by online platforms.

On the basis of these conclusions, the CMA recommended that the Government introduce legislation to enable the development of a new regulatory regime for the digital advertising market with the objectives of encouraging competition, promoting innovation and protecting consumers. With regard to the latter of these concerns, the CMA recommended that the regulator be given powers to introduce the following two key interventions:

- *The Choice Requirement Remedy (CRR)*: requiring platforms to give consumers the choice not to share their data for personalised advertising; and,
- *The 'Fairness by Design' duty*: placing a duty on platforms to ensure that they are maximising users' ability to make informed choices about the use of their personal data.

The present study has arisen out of Which?'s desire to understand how much value consumers place on having control over their data and the potential benefit from remedies in relation to consumer privacy protection.

## 1.2 Objectives

The objective for the research was to obtain user valuations for the CMA's proposed CRR via a stated preference methodology. This value evidence is intended to contribute to the broader evidence base in this area, with a view that it could potentially inform government decisions around potential legislation.

## 1.3 Report Structure

This is our final report for the study. It is organised as follows.

- Section 2 discusses the context for the study, including details of the proposed CRR and its motivation, and an overview of the literature surrounding the valuation of online privacy.
- Section 3 describes the survey methodology, including the stated preference design, questionnaire structure, pilot testing and main survey administration.

---

1 Online platforms and digital advertising (2020) CMA Market study final report, 1 July 2020.

- Section 4 presents findings on the characteristics of the population of Facebook and Google users, including service usage, awareness and attitudes towards targeted adverts and privacy settings.
- Section 5 reports our main findings, including descriptive and econometric analyses of the choices made by participants, and estimates of the willingness to pay values for the CMA's proposed CRR, disaggregated by segments of the sample.
- Section 6 concludes the study.

Appendix A contains the main survey questionnaire; Appendix B contains demographic characteristics of the sample data; Appendix C contains details and results from econometric modelling analyses, Appendix D presents value estimates of online privacy reported in the literature, and Appendix E contains a peer review of the present study by Prof. Ken Willis of Newcastle University.

## 2. Context for the Study

### 2.1 Policy Context

#### Regulatory landscape

The rise of the digital economy has impacted our society in a number of positive ways, including providing new means of communication and collaboration; new products that feature a strong service component; the role of data as a driver of economic growth; the automation of tasks with artificial intelligence and the emergence of new business models such as digital platforms (OECD, 2019). However, the rapid growth of digitisation has also created a number of regulatory challenges. For example, huge economies of scale and scope make it difficult to maintain competitive conditions (OECD, 2018).

An ongoing programme of work within the UK, and internationally, has been running with the aim of reviewing and developing regulatory structures suitable for modern digital markets. In the UK, this work includes a report on commercial use of consumer data;<sup>2</sup> research published on the effects of brand loyalty and online reviews on consumer behaviour, and how firms utilise search engine optimisation and paid searches to gain visibility;<sup>3</sup> research published on the use of pricing algorithms to facilitate collusion and personalised pricing;<sup>4</sup> and reforms proposed to the UK Government in relation to CMA's existing competition and consumer powers to ensure that they were fit for digital markets.<sup>5</sup>

In July 2019, the CMA published its first Digital Markets Strategy.<sup>6</sup> This set out the CMA's priorities for its work in digital markets. Some of the priority focus areas included using CMA's existing tools and capabilities to understand the opportunities and challenges associated with digital markets, adapting these tools to meet the challenges of the digital economy, providing support to the Government's call for new regulatory structures in digital markets and focussing on potential remedies for the digital markets (CMA, 2019).

In July 2020, the CMA published the results of a market study into Online Platforms and Digital Advertising. Specifically, the CMA investigated the markets for internet search, social media and digital advertising and the role of Google and Facebook within them. The study concluded that the government should introduce legislation to develop a new regulatory regime for the digital market with the objectives of fostering competition and innovation and protecting consumers.

---

2 CMA, Commercial Use of Consumer Data, (June 2015), <https://www.gov.uk/cma-cases/commercial-use-of-consumer-data>

3 CMA, Online search behaviour: literature review, (April 2017): <https://www.gov.uk/government/publications/online-search-behaviour-literature-review>

4 CMA, Pricing algorithms research, collusion and personalised pricing, (October 2018): <https://www.gov.uk/government/publications/pricing-algorithms-research-collusion-and-personalised-pricing>

5 Letter from Andrew Tyrie to the Secretary of State for Business, Energy and Industrial Strategy, (2019): <https://www.gov.uk/government/publications/letter-from-andrew-tyrie-to-the-secretary-of-state-for-business-energy-and-industrial-strategy>

6 A key immediate precedent of the CMA study was the Furman review 'Unlocking Digital Competition', commissioned by HM Treasury, which reported in 2019.

With regard to consumer protection, the CMA market study recommended that the regulator introduce two key interventions to provide consumers with genuine choice and control over the use of their personal data. These interventions, proposed to be initially placed on firms with strategic market status (SMS) including Google and Facebook, included the following:

- *The Choice Requirement Remedy (CRR)* which requires digital platforms to allow consumers the choice of not sharing their data for personalised advertising; and
- *The 'Fairness by Design' duty* which places a duty on digital platforms to ensure that they are maximising users' ability to make informed choices about the use of their personal data.

### Proposed Choice Requirement Remedy

Under the CMA's proposed CRR, online platforms would be required to give consumers the choice of a basic service without personalised advertising, and the regulator would have the power to influence the presentation of the choice including defaults. Initially, this obligation would only be placed on firms with strategic market status including Google and Facebook as these platforms are used by the majority of consumers and hold the largest amount of consumer data.

The main purpose of the CRR is to provide consumers with genuine choice and control over the use of their personal data. Although access to online platforms can benefit consumers in a number of ways (e.g. by providing them with services such as search, social media, information and entertainment which increase their overall wellbeing), a number of surveys have indicated that consumers are not only concerned about their lack of knowledge regarding how their personal data are collected and used by online platforms but also uncomfortable about how little control they seem to have over their data.

The market study by CMA (2020a) identified a number of negative effects that may arise from the lack of consumer control over their data:

- Consumers may be forced to use an online platform that they otherwise would not have used if they were better informed and had a greater ability to make choices about the use of their data.
- Consumers may be forced to share more data than they would otherwise choose to.
- Consumers may not receive a fair return for their personal data which is used to earn revenues by online platforms.
- Consumers may not benefit from increased competition in the digital market which could have offered them more choice and innovation.
- Consumers may have a lower level of trust generally in the internet and digital service providers.

The study also found that, even where consumers are given a choice over the use of their data, they tend to be unable to make an informed choice due to how the choice is structured and described. This is understood to be likely to limit consumer engagement and lead to an outcome where some consumers end up sharing more of their data than they would otherwise wish to.

CMA (2020a) provides a discussion of the key design issues that the regulator would need to consider when implementing the CRR. These include the manner in which the choice is presented, and the conditions attached by the platform related to personalised advertising to the choice.

The choice over personalised advertising under CRR may be presented in the form of a choice screen which consumers would have to engage with prior to using the platform. An alternative would be for platforms to impose a default option and have consumers access their settings to change the default. In the event of a default option being used, it is argued to be beneficial to consumers if the default is set to opt-out of personalised advertising.

This recommended default option is in line with the findings from recent qualitative research conducted by Which?.<sup>7</sup> This indicated that consumers preferred to be able to choose whether the advertising they saw was targeted and to opt-in to targeted advertising rather than opt-out.

Under the proposed CRR, consumers who choose not to share their personal data would continue to see generic contextual adverts and receive a core service from the online platform. However, these consumers could be provided with incentives to accept targeted personal advertising as this would benefit consumers and also help online platforms manage their revenue. These incentives could be in the form of special offers or other benefits that are deemed legitimate by the regulator.

This shift in the burden on platforms to encourage consumers to accept personalised advertising is expected to foster competition and may lead platforms to offer consumers a fair return for their personal data.

### Responses to the CMA (2020) market study

Facebook and Google responded to the findings and recommendations of the CMA (2020a) study. Although, Google did not specifically mention the CRR in its submission to the CMA, it noted the following:<sup>8</sup>

- Google's current privacy policies are customer-friendly and easily accessible.
- Google users can access their privacy settings page directly from all of its services.
- Both logged-in and logged-out Google users have the option to accept/reject personalised adverts across Google Search, YouTube and all other websites that partner with Google to show ads.

Facebook made a specific mention of the CRR in its submission to the CMA and raised several arguments against this remedy:

- Most consumers prefer relevant personalised advertising over untargeted advertising, so the CRR is likely to make the majority of consumers substantially worse off.
- Contextual advertising, allowed for all customers under the CRR, is also privacy-intrusive and is likely to worsen consumer experience.
- The implementation of the CRR is likely to have an adverse impact on advertisers and in particular SMEs who benefit from personalised advertising.
- The CRR is likely to reduce revenues obtained via personalised advertising which will disallow online platforms to fund valuable services which benefit users.

---

7 Which? (2020): Are you following me? – Consumer attitudes towards data collection methods for targeted advertising.

8 Online Platform and Digital Advertising: Comments on the Market Study Interim Report by Google: [https://assets.publishing.service.gov.uk/media/5e8c8290d3bf7f1fb7b91c2c/200212\\_Google\\_response\\_to\\_interim\\_report.pdf](https://assets.publishing.service.gov.uk/media/5e8c8290d3bf7f1fb7b91c2c/200212_Google_response_to_interim_report.pdf)

- The CRR is likely to have a disproportionate impact on social media platforms and other display advertisers which use personalised targeting relative to, for example, search engines for whom similar value can be captured through the search query (i.e. in a contextual rather than personalised manner).

The CMA, upon careful consideration of the above arguments, concluded that despite some potential costs of the CRR, overall it would prove to be an effective regulatory tool that would protect and maximise benefits for consumers.

Following the market study, the UK government asked the CMA to lead a Digital Markets Taskforce to provide advice to the government on how to design and implement a pro-competition regulatory regime for digital markets. To inform its work, the Taskforce invited views from a wide range of stakeholders including online platform operators and customers of digital platforms such as consumers and businesses that relied on digital platforms to sell their products and services (e.g. providers of apps and UK online marketplace retailers).

In response to this call by the Taskforce, Which? Submitted a report wherein it welcomed the findings of the CMA study, particularly those in relation to consumer privacy protection.<sup>9</sup>

The report of the Digital Markets Taskforce<sup>10</sup> recommended the establishment of a Digital Markets Unit (DMU), within the CMA, to administer a pro-competition regime for online platforms. The DMU was set up on a non-statutory basis in April 2021.

The Government intends to consult on proposals for the new pro-competition regime in 2021. Following these consultations, the DMU will be authorised to enact the required legislation. In the interim, the DMU will continue to work towards promoting greater competition and innovation in the digital markets and protecting consumers and businesses from unfair practices.<sup>11</sup>

The present study was intended to provide empirical evidence on how much customers value the proposed CRR with a view to informing future consultations.

## 2.2 Privacy Concerns

In recent years, Which? has produced or commissioned a number of policy research reports relating to consumer attitudes towards online data collection and usage. These include the following:

- Which? (2018) Consumers and their data: Research review. Policy research report, March 2018.
- Which? (2018) Control, alt or delete? Consumer research on attitudes to data collection and use. Policy research report, June 2018. (Includes BritainThinks report).
- Frontier economics (2018) Platforms, brokers and connected devices: what are the implications of the use of consumer data. An economic review prepared for Which? March 2018.
- Oxera (2018) Consumer data in online markets, prepared for Which?, June 2018.
- Which? (2020) Connecting the world to fraudsters.

---

9 Which? (2020) Submission to the Call for Information from the Digital Markets Taskforce, July 2020, p.5–7. Ore

10 CMA (2020) A new pro-competition regime for digital markets: Advice of the Digital Markets Taskforce, December, 2020.

11 See <https://www.gov.uk/government/collections/digital-markets-unit>

- Which? (2020) Are you following me? Consumers' attitudes towards data collection methods for targeted advertising. Policy research report, June 2020.
- Which? (2021) Are you still following me? Consumers' attitudes to data collection methods for targeted advertising. Policy research report, forthcoming September 2021.

The latter two are of particular relevance to the present study. Which? Conducted a qualitative study in March/April 2020 to investigate consumers' attitudes towards data collection methods used for the purpose of targeted advertising. The study consisted of 23 semi-structured depth interviews with Facebook users residing in London. The main objectives of the study were to understand:

- consumers' spontaneous/uninformed attitudes towards, and knowledge of, targeted advertising and the data collection methods underpinning it;
- consumers' informed attitudes towards, and knowledge of, first-party and third-party data collection methods; and
- the type of consent that they wanted for first-party and third-party data collection methods.

The results of the study which was published in June 2020 showed the following:

- The majority of survey participants preferred to receive targeted rather than generic adverts. However, they were largely unaware of the various data collection methods that informed such targeted advertising. Participants were specifically unaware about third-party data collection methods and proposed that this information be made more transparent by Facebook.
- Overall, third-party data collection methods were considered to be less acceptable in comparison to first-party data collection methods. Reasons cited for this included the perceived legitimacy of the data collection, privacy issues, and the relevance and proportionality of the data collection for targeted adverts.
- The majority of participants preferred to opt-in to, rather than opt-out of, data collection for targeted advertising. Also, participants preferred more control over the type of data collection method used and hence preferred to be asked to consent to each data collection method individually.

Following this study, Which? Conducted a quantitative survey in June/July 2020 to further investigate the findings from its qualitative study. The quantitative study comprised of an online survey of 1,729 adult Facebook users living in the UK.

The results of the quantitative study supported and validated the findings of the qualitative study and concluded the following:

- Although awareness of first-party and third-party data collection methods was high, most participants were unaware regarding the extent of such data collection for purposes of targeted advertising.
- Once informed of third-party data collection methods, most participants felt that they had not given their informed consent to such data collection methods.
- Once informed of the data collection methods used to inform targeted advertising, some participants who had initially chosen targeted adverts switched to generic adverts.

- Acceptability of third-party data collection methods such as online tracking and customer lists was much higher among those participants who felt that they had given their informed consent to such data collection methods.
- The majority of participants preferred to opt-in to receive targeted adverts rather than having to opt-out so as to not receive them.
- Most participants preferred to be given the choice of opting into each of the data collection methods individually so as to provide them with greater control over the way in which their personal data was used.
- The majority of participants felt that Facebook was not particularly transparent with users around its third-party data collection methods. Those who felt that Facebook was transparent, were more likely to feel that they had given informed consent and that third-party data collection methods were acceptable.

Based on the above research and analysis, Which? Supported the recommendations proposed in the CMA market study. With regard to the ‘choice requirement’, Which? has argued for a default opt out position with respect to digital advertising on the basis that this maximises consumer protection and is the preferred choice of customers. Which? has also recommended extending the remedy to allow consumers to choose the sources from which their data are collected rather than solely whether or not they want to receive personalised advertising.

## 2.3 Valuations of Privacy

Although no previous studies have directly sought to value the proposed CRR, there have been a number of academic studies focused on valuing privacy and personal data. Empirical research in this area has focused principally on two issues: the apparent dichotomy between privacy attitudes and actual privacy behaviour of consumers, and estimation of the value that consumers assign to privacy and personal data.

Several national surveys on consumer attitudes towards use of personal data, (e.g. Turrow et.al 2009 and Rainie et.al 2013 conducted in the US and Statista, 2021 conducted in the UK) have shown that online privacy is a significant concern among internet users. However, actual internet usage of consumers show that the vast majority of these consumers continue to use information technologies that track and share their personal information with third parties. This phenomenon has been termed as the “privacy paradox”.

A number of studies have attempted to understand the reasons behind the privacy paradox (e.g. Berendt, Gunther and Spiekermann 2005; Norberg, Horne and Horne 2007). These studies have identified a number of co-existing explanatory factors, including asymmetric information, bounded rationality and various decision-making heuristics. For example, some individuals may be genuinely unaware of the extent to which their personal data is being compromised (McDonald and Cranor, 2010; Hoofnagle and Urban, 2014), while some others may be unaware of the availability of various solutions to their privacy concerns such as privacy-enhancing technologies. Further, even if individuals are well-informed about online privacy settings, their privacy behaviour may be impacted by behavioural biases such as immediate gratification or status quo bias (John, Acquisti and Loewenstein, 2011).

Studies have also highlighted a particular challenge involved in valuing privacy: whether privacy should be valued as the lowest price that an individual is willing to accept to reveal

his/her personal data which he/she initially owned (WTA) or the maximum price that an individual would be willing to pay to protect his/her personal data which he/she did not initially own (WTP). For private goods, these issues can be resolved via a market for the good which reflects the reservation prices of buyers and sellers in the market and hence accurately captures the price of goods. However, no such market exists for personal data. Although personal data is traded among firms on a continuous basis, consumers do not have access to (and frequently remain unaware of) such markets where they can offer to share or buy back their data at a price. Further, a lack of consumer awareness of privacy trade-offs make it difficult for market outcomes to accurately capture individuals' true valuations of privacy (Berthold and Bohme 2010). Hence privacy can be considered to be a public good that is prone to market failure (Saetra,2020).

Due to the aforementioned hurdles, it has been a challenge to obtain reliable estimates of privacy. Some empirical studies have set up experiments in which participants were presented with scenarios involving actual privacy and financial trade-offs, while others have conducted hypothetical surveys to quantify the values that consumers assign to their privacy or personal data.

Studies involving surveys/experiments to estimate individual's willingness to accept payment for the sharing of their personal data include Huberman, Adar and Fine (2005), Wathieu and Friedman (2007), Cvrcek et.al (2006), Hui, Teo and Lee (2007), Chellappa and Sin (2005), Tedeschi (2002), and Spiekermann, Grossklags and Berendt (2001). Studies involving surveys/experiments to estimate individual's willingness to pay for protecting their personal data include Rose (2005), Tsai et. Al (2011), Jentzsch, Preibusch, and Harasser (2012), Beresford, Kubler, and Preibusch (2012), Varian, Wallenberg, and Woroch (2005) and Png (2007). All of these studies analysed the value of personal information from the users' perspective. Most have shown that privacy valuations are highly context dependent.

In contrast, Olejnik, Minh-Dung and Castelluccia (2014) attempted to quantify the value of personal data from the advertisers' perspective. The study analysed and quantified the impact of Real-Time Bidding (RTB) and Cookie Matching (CM) technologies on user' privacy. The study found that RTB creates a market where elements of users' browsing histories are sold at auctions to internet advertising companies for amounts lower than \$0.0005 per person. However, these prices varied depending on factors as the visiting site, time, user's physical location and user's profile.

Empirical studies related to the analysis and quantification of the value of online privacy are rare. Some examples include Hann et.al (2007), Savage and Waldman (2013) and Prince and Wallsten (2020). Hann et.al (2007), a seminal contribution in this area, used conjoint analysis in which each participant (originating from the US and Singapore) was asked to assess trade-offs between privacy protection policies of organisations (i.e. collection, error, unauthorised secondary use, and improper access to personal information online) and/or promotions and/or convenience. The results of the study revealed that for US subjects, the value of protection of personal information was between \$30.49 and \$44.62 while for Singapore subjects the value was estimated at S\$57.11.

Savage and Waldman (2013) conducted a choice experiment with US smartphone consumers in which each participant was presented with a choice set containing an app that was currently traded in the marketplace and five new apps that had similar functionality to the market app but varied in terms of their price, advertising and privacy permission levels. The results of the study showed that a representative consumer was willing to make a one-time payment of \$2.28 to conceal their browser history, \$4.05 to conceal their contacts list, \$1.19 to conceal their location

data, \$1.75 to conceal their phone's identification number, \$3.58 to conceal the contents of their text messages and \$2.12 to eliminate advertising.

Prince and Wallsten (2020) used discrete choice surveys to estimate individuals' willingness to accept to give up various forms of online privacy. The study considered participants across various countries (US, Mexico, Brazil, Colombia, Argentina and Germany) to measure values for a range of data privacy types (personal information on finances, biometrics, location, networks, communications, and web browsing). The results of the study showed that on average individuals were willing to accept the following payments from platforms to allow them to share their information with third parties: \$8.44/month to share their bank balance, \$7.56/month to share their fingerprint information, \$6.05/month to access and read their texts, \$5.80/month to share information on their cash withdrawals, \$1.82/month to share their location data and negligible amounts to be sent adverts via SMS.

Overall, the empirical research indicates that consumers may not always adopt a rational utility-maximising approach to decide on their online privacy. In fact, other approaches that account for factors such as incomplete and asymmetric information, context dependency and behavioural and cognitive biases in decision-making may better reflect individuals' choices regarding their online privacy. Hence, the focus has shifted towards using behavioural economics research to understand the factors influencing online privacy values.

The behavioural economics literature discusses a number of behavioural biases as well as the importance of choice architecture that may influence consumers' privacy valuations. Some of the behavioural biases that can have an impact on consumers' online privacy decisions include: status quo bias, framing or presentation effects, anchoring effects, loss aversion/endowment effects, hyperbolic discounting effects, inattention and overconfidence effects (CMA, 2020b). These are explained below.

### Behavioural Biases

- *Status quo bias*: this occurs when consumers prefer the current state and consider any change to be a loss. Studies such as Goldstein (2008), Lai and Hui (2004), Leon et.al (2011) and Acquisti and Gross (2006) have shown that in an online environment, consumers were likely to choose the default options/settings. This was because most consumers assumed that the default configurations presented to them were designed by companies to protect them.
- *Framing or presentation effects*: this occurs when the manner in which control over disclosure of personal information is presented to users have an influence on their privacy decision-making. Studies such as Brandimarte et.al (2012) and Adjerid et.al (2013) have conducted experiments to show that stricter privacy controls/notices presented to users were likely to lead them to disclose more personal information and vice-versa.

Acquisti et.al (2015) explored how the framing and presentation effects might 'nudge' users to reveal their personal information. For example, the study showed that in promotional emails, the option to unsubscribe from the mailing list was placed in small text and bland colours at the bottom of the email. The study argued that this was done to discourage users from choosing to protect their privacy.

- *Anchoring effects*: this occurs when users' decision-making is affected by a particular reference point or anchor i.e. users rely heavily on the very first piece of information they learn, which has a significant impact on their final decision. For example, Acquisti et.al (2012) showed that the order effects played an important role in influencing users' privacy decision-making. The study found that disclosure of personal information was anchored by the initial questions

in a survey i.e. participants tended to divulge more sensitive personal information when the initial questions in the survey were more privacy-intrusive than the subsequent questions.

- *Loss aversion effects*: this occurs when individuals prefer avoiding losses compared to making equivalent gains. Studies such as Grossklags and Acquisti (2007) and Winegar and Sunstein (2019) have conducted surveys and experiments to show that consumers were willing to accept a much higher payment to allow access to their personal data in comparison to the amount that they were willing to pay to maintain their data privacy.
- *Hyperbolic discounting effects*: this occurs when individuals choose smaller but immediate rewards rather than larger but future rewards. Studies such as Acquisti (2004) and Acquisti and Grossklags (2005) have shown that the privacy paradox might exist due to such discounting effects whereby users heavily discount high future risks associated with disclosing their personal information (e.g. identity thefts) in the current time period.
- *Inattention and Optimism/Overconfidence effects*: Acquisti et.al (2017) argue that inattention effects whereby users focus their attention only to a part of the information or options available to them can influence their decision-making in an online setting. Similarly, users' privacy decision-making may be impacted if they underestimate the privacy risks that they face or overestimate their ability to assess these risks accurately.

### Choice Architecture

Johnson et.al (2012) argue that how the choice is structured and how the choice is described can have a significant impact on users' decision-making. Influential factors include the number of options presented, the use of default options/settings, the use of decision aids (e.g. recommendations), and the way in which different options are presented.

Acquisti et.al (2015), and reports published by the Norwegian Consumer Council (2018), study the effect of choice architecture on users' decision-making in the online environment. Acquisti et.al discuss an example of a sign-up procedure to an online service that presented users with the default option of keeping their profile public. In addition, the 'keep profile public' option was presented in a brighter colour and situated on the right-hand side of the dialog box where typically options to move the user to the next stage of the sign-up process was situated. Similarly, the NCC reports analysed the interfaces used by Facebook and Google and found that, in most cases, privacy-intrusive settings were set as the default option. Similar effects of the choice architecture on users' choices were found in the case of mobile app users (see Egelman et.al (2013) and Zafeiropoulou et.al (2013)).

More recent research in this area is focussed on studying how firms exploit behavioural biases to design choice architecture in order to encourage users to opt for unintended privacy-intrusive options. These practices known as 'sludge techniques' are analysed in studies such as Shklovski et.al (2014), Quinn (2016), Mathur et.al (2019) etc. However, some other studies (Kelley et al (2010), Acquisti et al (2017) etc.) have shown that the choice architecture can also be designed to improve consumer decision-making.

Given the potential challenges associated with valuation of online privacy detailed above, a thorough and careful stage of development work was important within our study to fully test and assure the validity of the survey instrument. Details regarding the development phase and the methodology used in our study are discussed in the next section.

## 3. Methodology

### 3.1 Introduction

The core objective for the survey design was to understand and estimate the value of the CMA's proposed CRR. Section 3.2 describes the approach taken in the design of the survey to obtain these estimates. Section 3.3 details the outline of the questionnaire in full, including the additional questions that were included, Section 3.4 describes the detailed development and testing programme carried out to ensure the validity of the main survey instrument, Section 3.5 discusses the main survey administration and Section 3.6 presents diagnostics concerning the sample quality, and the sensitivity samples examined in the research.

### 3.2 Contingent Valuation Design

The nature of the CRR policy intervention differs in an important way from the type of intervention usually valued by stated preference valuation methodologies. Typically, a policy, programme or project is valued by asking participants to make a referendum choice over whether they would prefer to have the intervention, and its associated benefits, but pay a price in terms of higher taxes or bills, or to not have the intervention and not pay the price. In the case of the present study, this construct was considered inappropriate due to the lack of a plausible payment vehicle for the CRR as a policy. It did not seem plausible, for example, that taxes would need to rise to pay for the CRR.

The valuation approach adopted was hence instead based on the assumption that the CRR would effectively add a new option to the choice set for users. The core valuation framework was thus based around offering survey participants a choice between two alternatives:

- Full sharing of personal data / targeted adverts; and
- No sharing of personal data / generic adverts.

Despite these options not being new to Google and Facebook users, the vast majority of Google and Facebook users were known not to amend their privacy settings to prevent personalised advertising despite having the option of doing so. Indeed, the CMA concluded in its market study that this is due to there being many barriers to engagement, relating to the choice architecture and the length and complexity of terms and conditions.

Where users have reviewed their settings, the valuation approach assumes that the CRR creates no additional value. This is on the basis that it cannot be considered to create a new alternative for those that have already exercised choice between existing alternatives. This is a conservative assumption as there is likely to be some additional value even to these users of having the choice made more transparent and accessible. This value is not counted in the analysis.

Under the standard assumptions of the logit model of discrete choice, the change in expected consumer surplus arising as a result of a change in the choice set is given by the following formula:

$$\Delta E(CS_i) = \frac{1}{\alpha_i} \left[ \ln \left( \sum_{j=1}^J e^{V_{ij}} \right) - \ln \left( \sum_{k=1}^K e^{V_{ik}} \right) \right] \quad (1)$$

Where  $CS_i$  is the consumer surplus for the  $i^{\text{th}}$  consumer,  $\alpha_i$  is the marginal utility of money for the  $i^{\text{th}}$  consumer,  $j = \{1, \dots, J\}$  and  $k = \{1, \dots, K\}$  are the alternatives post and pre-intervention and  $V_{ij}$  and  $V_{ik}$  are the deterministic components of utility for the post and pre-intervention for the corresponding alternative and consumer.

Application of this approach requires the measurement of  $\alpha$ , the marginal utility of money, which accordingly is measured in the design via inclusion of a financial incentive applied to the choice of targeted adverts.

Two alternative payment vehicles were included in the design:

- A subscription fee associated with using the ‘No sharing of personal data’ option i.e. using the service (Facebook/Google) with generic adverts;
- A reward (e.g. PayPal credit and Amazon vouchers/gift card) associated with the ‘Full sharing of personal data’ option i.e. using the service and receiving targeted adverts.

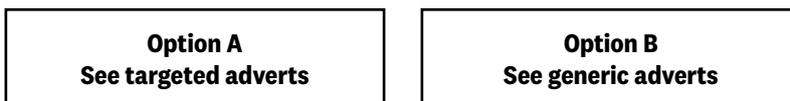
There were two key reasons for testing both forms of payment vehicle. Firstly, both were considered credible alternatives for how the CRR may be implemented in practice. Secondly, there is widespread evidence that people are less sensitive, per pound, to money gains, like rewards, than they are to money losses, like the subscription fee. Applying both types of payment vehicle allowed for a fuller assessment of the value of the CRR to consumers than applying just one in isolation.

Earlier research, including that by Which?, had shown that consumers tend to be unaware of the extent to which online platforms such as Google and Facebook collect personal data from them, including from off-platform online activity. In order to test the effect of such information on the choice of the alternatives, the survey was designed to include: first, an uninformed choice exercise; then, questions to establish the participant’s baseline levels of awareness; then, a presentation of information about Facebook/Google’s use of personal data; and then, a repeated choice version of the choice exercise.

In addition to the core valuation exercise, the survey included a further choice question offering multiple alternatives that differed according to whether the data was collected via first-party methods (ie. data is collected from platforms owned by the organisation) or third-party methods (data is collected via tracking on third party websites and apps and third-party customer lists). This question was included to allow for exploration of attitudes associated with the sharing of different types of privacy protection.

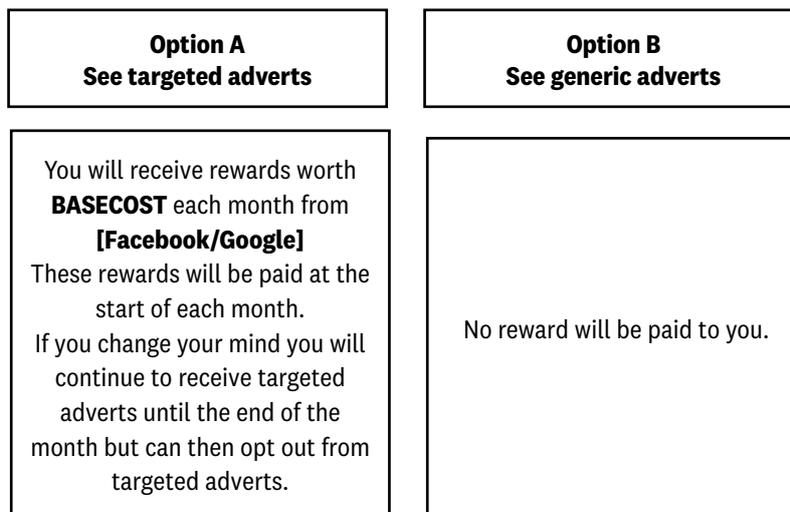
Figures 1–3 show examples of the type of choice questions asked under the uninformed scenario. Initially participants were asked to choose between the two alternatives of seeing targeted adverts and seeing generic adverts without any mention of a reward or payment (see Figure 1). The order of the two options was randomised across the sample, with half the sample seeing Option A and Option B as shown in the figure below, while the remaining half saw Option A as “See generic adverts” and Option B as “See targeted adverts”.

**Figure 1:** Uninformed Choice: No Reward or Payment

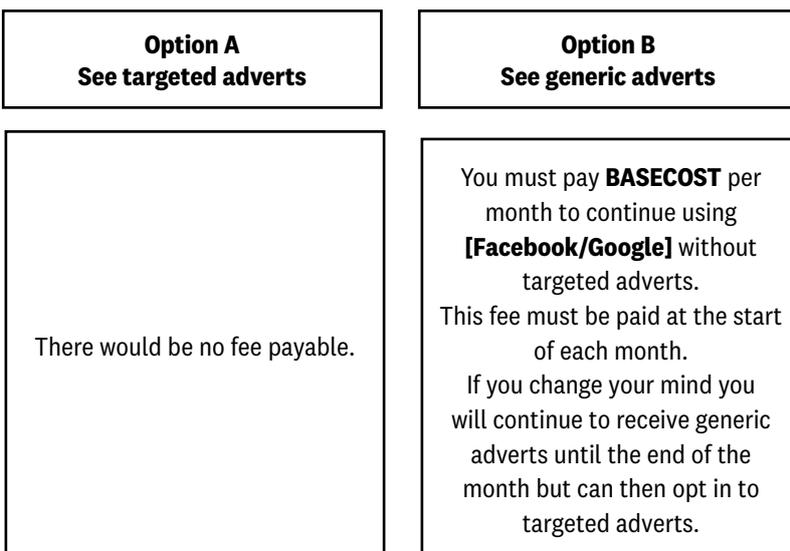


Next, a similar choice was presented to participants but with a financial incentive attached. Participants were assigned at random either to the Reward treatment, in which case they saw a question like Figure 2; or, they were assigned to the Fee treatment, and were shown a question like Figure 3. (The assignment was fixed for the duration of the survey; i.e. participants did not switch between Reward and Fee questions during the survey.)

**Figure 2:** Uninformed Choice (Binary): Reward



**Figure 3:** Uninformed Choice (Binary): Payment



The reward and fee amounts (BASECOST per month) were selected at random from the list = {£0.50, £1.00, £2.00, £3.00, £5.00}.<sup>12</sup> Depending on their initial response, they were then shown the same choice cards but with the amount doubled if they chose “See generic adverts” in the initial exercise and halved if they chose “See targeted adverts” in the initial question.

This approach is known as the double-bounded dichotomous choice (DBDC) contingent valuation (CV) format. In comparison to open-ended and payment card questions, participants find it easier to understand the format, are much less likely to give protest responses, and have less (no) incentive to misrepresent their preferences (Carson and Groves, 2007). This format therefore minimises non-responses and avoids outliers.

The double-bounded version of the dichotomous choice question provides greater statistical precision than the single question version due to the additional information provided by the follow-up question. However, the response to the second question can be influenced by the amount shown in the first question and cannot therefore be considered an independent observation. This dependence must be considered in the analysis of the data to avoid obtaining a biased value estimate of the value. (See Appendix C for details of how the econometric modelling dealt with this issue.)

Figures 4–6 show examples of the type of choice questions asked under the informed scenario. First, participants were presented with information about the ways in which Facebook/Google collect their data i.e. first party data collection, third party online tracking and third-party customer lists. Next, they were asked to choose between two alternatives: Full data sharing vs. No data sharing without any mention of a reward or payment (see Figure 4). Again, the order of the two options were presented to the participants at random i.e. some participants saw Option A and Option B as shown in the figure below, while others saw Option A as “No data sharing” and Option B as “Full data sharing”.

**Figure 4:** Informed Choice (Binary): No Reward or Payment

	Option A: Full data sharing	Option B: No data sharing
Type of adverts shown	Targeted	Generic
<b>Where information about you is collected from:</b>		
<b>1st party data collection:</b> Instagram/ Facebook activity i.e. which posts you like, content you watch or engage with, places you ‘check-in’ to and your profile data	Yes	No
<b>3rd party online tracking:</b> information from other 3rd party organisations you use that have a relationship with Facebook	Yes	No
<b>3rd party customer lists:</b> matching your profile with customer lists uploaded by other organisations you have used/signed up to	Yes	No

In early versions of this choice, and subsequent informed choice questions, the ‘YES’ and ‘NO’ labels were substituted with green ticks and red crosses. However, in the course of pilot testing, as discussed in Section 3.4, there was some evidence that this format was leading participants into choosing the ‘Targeted’ choice. To address the issue, the ticks and crosses were replaced with the format shown above.

12 Payment levels were explored in a series of in-depth interviews and tested at the pilot stage (see section 3.4).

The next choice questions (Figure 5 and Figure 6) asked participants to choose between ‘Targeted’ and ‘Generic’ with a reward or payment incentive to choose ‘Targeted’. The monthly reward and payment amounts were the same as those used for the uninformed question and, again, depending on their initial response, they were shown the same choice cards but with base cost doubled if they chose “See generic adverts” in the initial question and base cost halved if they chose “See targeted adverts” in the initial question.

**Figure 5:** Informed Choice (Binary): Reward

	Option A: Full data sharing	Option B: No data sharing
Type of adverts shown	Targeted	Generic
<b>Where information about you is collected from:</b>		
<b>1st party data collection:</b> Instagram/ Facebook activity i.e. which posts you like, content you watch or engage with, places you ‘check-in’ to and your profile data	Yes	No
<b>3rd party online tracking:</b> information from other 3rd party organisations you use that have a relationship with Facebook	Yes	No
<b>3rd party customer lists:</b> matching your profile with customer lists uploaded by other organisations you have used/signed up to	Yes	No
<b>Value of reward paid to you each month</b>	<b>BASECOST</b>	<b>£0</b>

**Figure 6:** Informed Choice (Binary): Payment

	Option A: Full data sharing	Option B: No data sharing
Type of adverts shown	Targeted	Generic
<b>Where information about you is collected from:</b>		
<b>1st party data collection:</b> Instagram/ Facebook activity i.e. which posts you like, content you watch or engage with, places you ‘check-in’ to and your profile data	Yes	No
<b>3rd party online tracking:</b> information from other 3rd party organisations you use that have a relationship with Facebook	Yes	No
<b>3rd party customer lists:</b> matching your profile with customer lists uploaded by other organisations you have used/signed up to	Yes	No
<b>Monthly subscription fee</b>	<b>£0</b>	<b>BASECOST</b>

Finally, in order to estimate the values of different types of privacy protection, participants were presented with a multi-alternative choice question. The choice question contained five options of varying levels of privacy protection and the associated rewards/payments. Figure 7 and Figure 8 shows examples of the multi-alternative choice question with options of monthly rewards and subscription fees respectively.

The amounts shown for each alternative were specified as fixed proportions of the BASECOST amount, which itself was fixed across the survey for each participant. This approach was taken to enable direct comparison of choices amongst alternatives across sample segments whilst still retaining consistency with earlier questions regarding the cost amount associated with the fully

targeted and fully generic options, and the associated variation in overall cost levels implied by this. This approach meant that the relative difference in cost amounts across alternatives was constant for all participants, but the absolute difference varied according to the BASECOST amount assigned.

**Figure 7: Informed Choice (Full): Reward**

	Option A Full data sharing	Option B 1st party data and 3rd party customer lists	Option C 1st party data and 3rd party online tracking	Option D 1st party data only	Option E No data sharing
<b>Type of adverts shown</b>	Targeted	Targeted	Targeted	Targeted	Generic
<b>Where information about you is collected from:</b>					
<b>1st party data collection:</b> Instagram/ Facebook activity i.e. which posts you like, content you watch or engage with, places you 'check-in' to and your profile data	Yes	Yes	Yes	Yes	No
<b>3rd party online tracking:</b> information from other 3rd party organisations you use that have a relationship with Facebook	Yes	No	Yes	No	No
<b>3rd party customer lists:</b> matching your profile with customer lists uploaded by other organisations you have used/signed up to	Yes	Yes	No	No	No
<b>Value of reward paid to you each month</b>	BASECOST	0.75 *BASECOST	0.75 *BASECOST	0.5 *BASECOST	£0

**Figure 8: Informed Choice (Full): Payment**

	Option A Full data sharing	Option B 1st party data and 3rd party customer lists	Option C 1st party data and 3rd party online tracking	Option D 1st party data only	Option E No data sharing
<b>Type of adverts shown</b>	Targeted	Targeted	Targeted	Targeted	Generic
<b>Where information about you is collected from:</b>					
<b>1st party data collection:</b> Instagram/ Facebook activity i.e. which posts you like, content you watch or engage with, places you 'check-in' to and your profile data	Yes	Yes	Yes	Yes	No
<b>3rd party online tracking:</b> information from other 3rd party organisations you use that have a relationship with Facebook	Yes	No	Yes	No	No
<b>3rd party customer lists:</b> matching your profile with customer lists uploaded by other organisations you have used/signed up to	Yes	Yes	No	No	No
<b>Monthly subscription fee</b>	£0	0.25 *BASECOST	0.25 *BASECOST	0.5 *BASECOST	BASECOST

### 3.3 Questionnaire Structure

The full survey questionnaire was structured as follows:

- Participant screening to ensure a nationally representative sample of UK adults 18+ (quotas were set on age, gender, region and approximate SEG) at the start of the survey
- Usage of Google and Facebook services
  - Participants were then randomly allocated between Facebook and Google unless they only used one of the platforms i.e. they used the other less than once a month or never. Participants who selected ‘never use this service’ or use less than once a month for all services screened out of the survey at this stage. For the purposes of this research, using only Instagram did not qualify a participant as a Facebook user
- Device and operating system used
- Targeted advertising awareness and preferences for targeted/generic adverts without a financial incentive/fee attached
- Uniformed choice section
- Information about the three ways in which data is collected by Facebook/Google i.e. first party data collection, third party data collection and third-party customer lists
- Informed choice section
- Final classification questions about the respondent

### 3.4 Development and Testing

Several phases of development and testing research were undertaken before the main survey was implemented. First, a preliminary phase of qualitative research was undertaken comprising 10 in-depth interviews with Google and Facebook users in January 2021. The purpose of this phase was to explore initial designs for the main choice questions and test the credibility of reward and payment versions, payment vehicle options, and reward and payment levels. The interviews were conducted remotely by Zoom and lasted between 45 and 60 minutes.

This research found that, in general, the main choice exercise appeared to work well although some of the scenarios were felt to be more plausible than others. As would be expected, a reward was immediately seen as more attractive than a subscription, and for most, was also seen as more realistic. However, both options were accepted as plausible.

There were some useful indications from the discussions around what payment vehicles would be plausible, and around the levels of reward and payment that could be used in the choice questions. These were used to set initial levels for the cognitive interviews and pilot.

A series of 10 cognitive interviews were then undertaken in March 2021 to test the full programmed quantitative survey. One issue raised during these interviews was a confusion over the question on settings and it was recommended that screenshots of the relevant settings page were added to the survey. A small number of minor changes were made to take into the pilot.

A pilot survey was conducted online in the last week of April 2021 using sample sourced from an online panel. A total of 396 fully completed interviews were achieved for the pilot survey. The average completion time was 7 minutes 13 seconds. Once the pilot sample was achieved, the data was subject to rigorous quality checks.

The findings from the pilot analysis are as follows:

- There were no issues with recruitment, and the rate at which interviews were added to the dataset was as expected.
- The length of survey was as expected.
- There were no routing errors. However, to simplify the analysis and allow for richer data we recommended that all participants answer all contingent valuation questions rather than routing away those that chose targeted adverts when there was no reward or fee.
- Analysis of the reasons for choices showed that these were valid in the vast majority of cases. However, some reasons raised a concern that the informed choice without payment vehicle was being influenced by the payment vehicle. Hence, we recommended a note to encourage participants to respond to the question as intended at this point. Furthermore, there was a concern that participants may have found the use of green ticks and red crosses to represent whether or not particular types of data were being shared to be leading. Hence, we recommended revising these to YES and NO respectively.
- The range of cost levels included in the choice questions seemed appropriate, with the vast majority of the sample choosing 'Targeted adverts' at the maximum reward and fee levels.
- The results from the main binary choices between generic and targeted adverts broadly conformed to expectation, in the sense that cost sensitivity was as expected, and those offered rewards were less sensitive to the size of the reward than those shown fees were sensitive to the size of the fee.
- Simple econometric models were estimated and performed reasonably well. However, valuations were higher than expected. We anticipated more realistic valuations being obtained given a more thorough econometric modelling investigation than was feasible at the pilot stage.
- Results from the final exercise on multiple data sharing levels were consistent with binary choice results at the aggregate level, in respect of the proportions choosing full data sharing vs no data sharing. However, patterns of choices did not vary with cost in the manner expected. We recommended revising the ticks and crosses to be consistent with our recommendation for the informed binary choice question.

### 3.5 Main Survey Administration

The main survey was conducted online from 19 May to 1 June 2021. The sample was sourced from an online panel.

8,273 entered the survey, 4,022 participants were screened out of the survey at the scoping stage and 237 participants qualified but dropped out of the survey part way through. A total of 4,014 fully completed interviews were achieved. The average completion time was 7 minutes 16 seconds.

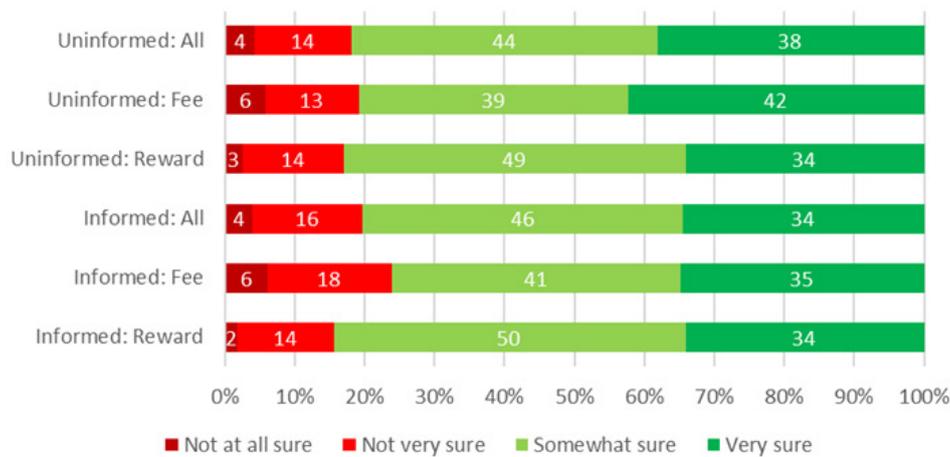
A note of caution is due with regard to the representativeness of the sample. Although the sample was designed to be representative by selecting Google and Facebook users at random from a Census-based sampling frame, the sampling frame itself was comprised of people who had signed up to an online panel. One might expect that online panel members, on account of them having signed up to complete surveys on a regular basis, are likely to have a greater willingness to share data about themselves than the general population of users. If this is the case, the valuations obtained may prove to be an underestimate of the true population valuations. Accordingly, one may consider the reported valuations as being conservative estimates of the true values of the CRR to Google and Facebook users.

### 3.6 Data Diagnostics and Sensitivity Samples

#### Confidence in contingent valuation choices

Participants were asked to state how sure they were that they would make the same choice in practice as they did in the survey if the choice was really offered to them. This ‘confidence question’ was asked after the second contingent valuation question (i.e., the first contingent valuation question which included a fee/reward) in both the uninformed and informed set of choices. As shown in Figure 9, around 80% were ‘Very sure’ or ‘Somewhat sure’ that they would make the same choices in a real choice situation.

**Figure 9:** Participants’ confidence in their contingent valuation choices



Base: 2,000 (Fee), 2,014 (Reward)

The proportion of participants who were sure (very or somewhat) was high across information treatments (uninformed/informed) and payment vehicles (fee/reward). Confidence is statistically lower (at 5%) for informed choice than for uninformed choice, but the difference is practically small. The proportion of participants who were ‘Very sure’ or ‘Somewhat sure’ of their choices was statistically lower (at the 1% level) for informed than for uninformed choice, but the difference was negligible (80.3% vs. 81.9%).

Regardless of whether choices were uninformed or informed, participants who were required to pay a fee to avoid receiving targeted adverts tended to have lower confidence than those who were offered a reward for choosing to receive targeted adverts. For informed choices, the proportion of those who were ‘Very sure’ or ‘Somewhat sure’ was around 8 percentage points lower in the fee treatment than in the reward treatment, the difference being highly statistically significant.

Stated confidence levels were fairly strongly correlated between uninformed and informed choices (Spearman rank correlation of 0.63<sup>13</sup>). In fact, 90% of those who were sure (‘Very sure’ or ‘Somewhat sure’) of their uninformed choices were also sure of their informed choices.

There is a concern amongst some authors in the non-market valuation literature that willingness to pay estimates based on hypothetical choices may overestimate actual willingness to pay, i.e., willingness to pay in situations which involve actual payments as opposed to hypothetical ones. Several studies across a wide range of valued goods have found a ‘hypothetical bias’ (see Ready et al. 2010 for a review). Hypothetical bias has been found to be inversely

13 Correlation is measured on a 0-1 scale. A value of 1 implies perfect rank correlation.

correlated with stated levels of confidence. Therefore, we examine, within the econometric analysis, how choices made by participants who were not sure of their choices impact our main estimates.

### Option order effects

The order of response options was randomised in the contingent valuation questions, as explained in Section 3.2.<sup>14</sup> This helps to mitigate potential bias in contingent valuation choice proportions due to option order effects, such as a greater tendency for participants to pick the first option on the left rather than the first on the right. We found that left-to-right bias was statistically significant at the 5% level (or close to being significant) in the dichotomous contingent valuation questions which did not include a payment vehicle (reward or fee) and in the final choice question which included multiple data-sharing options.<sup>15</sup> However, the effects were practically small. For example, 44% chose ‘Targeted adverts’ when it was shown as option A compared to 41% when shown as option B in the first of the uninformed choice questions. Random assignment of response option orders mitigates bias as the estimated proportion choosing ‘Targeted adverts’ is an average of the proportions under each ordering treatment.

### Completion time

While the average time taken to complete the questionnaire was 7 minutes and 16 seconds, around 5% of participants completed the questionnaire in under 3 minutes as shown in Table 2. There is a possibility that ‘speeders’ do not consider the options carefully enough or may even respond randomly to complete quickly. We therefore test the sensitivity of key findings to excluding participants whose completion time was less than the 10<sup>th</sup> percentile of the distribution of completion times.

**Table 2:** Distribution of completion times

Proportion of participants	Maximum completion time
1%	158 seconds
5%	182 seconds
10%	205 seconds
25%	263 seconds
50%	355 seconds
75%	494 seconds
90%	720 seconds
95%	919 seconds
99%	1748 seconds

Base: 4,014

<sup>14</sup> The order was randomised across participants, but not across questions, i.e., any given participant saw response options in the same order across questions to mitigate response error.

<sup>15</sup> Pearson chi-square tests of independence between contingent valuation answer and option order.

## 4. Usage, Awareness and Attitudes

### 4.1 Introduction

This section presents findings on the characteristics of the sample of Facebook and Google users in regard to:

- usage of Facebook and Google services, including the device used to access these services (see Section 4.2)
- awareness of, and attitudes towards, targeted adverts (see Section 4.3)
- privacy settings (see Section 4.4)
- attitudes towards data sharing (see Section 4.5).

The descriptive analysis in this section sets the stage for our analysis of contingent valuation (CV) choices in section 5 and provides valuable insights for interpreting our findings from analysis of CV choices. Comparisons of sample characteristics against findings from an earlier Which? survey provide a means of assessing the representativeness of the sample of internet users on which our main findings are based.

### 4.2 Service and Device Usage

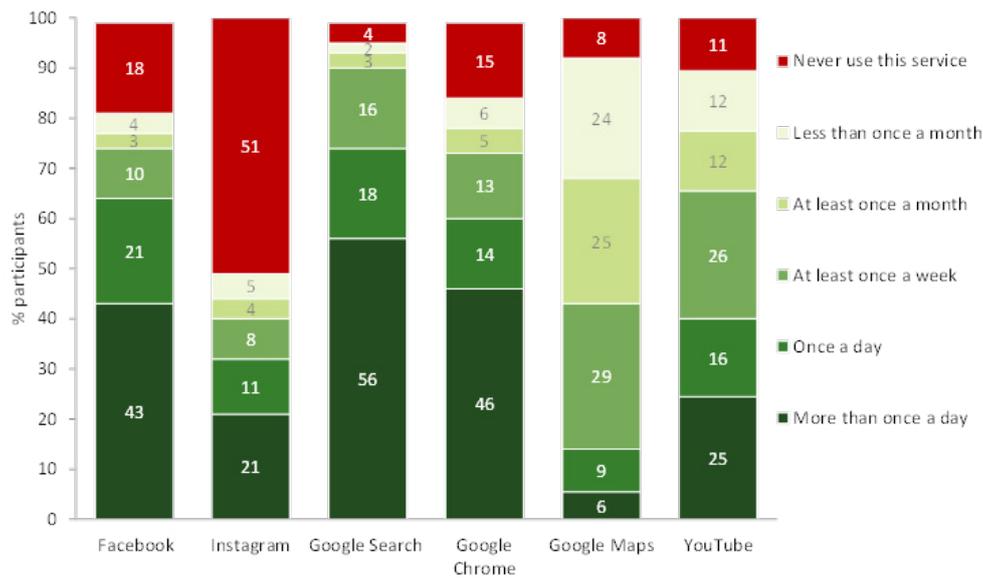
Participants were asked how often they used each of the following in the previous three months:

- Facebook (the main app rather than Messenger)
- Instagram
- Google Search – a search engine that allows you to search the web (similar to Bing and Yahoo)
- Google Chrome – a web browser (similar to Firefox, Internet Explorer, Safari)Google Maps
- YouTube.

For the overall sample Google Search was used most (over three quarters used it at least daily: 56% more than once a day and 18% once a day). Facebook was used at least once a day by 64% (43% more than once a day and 21% once a day). Google Chrome was the third most used service (used at least once a day by 60% (46% more than once a day and 14% once a day)).

YouTube was used at least once a day by 41% of the sample. Instagram was used least with over a half of the sample not using it at all. Only 8% never used Google Maps but it was used the least frequently of all services. See Figure 10.

**Figure 10:** Use of Facebook and Google services in the preceding three months



Base: 4,014

The main statistically significant differences<sup>16</sup> for the samples using Facebook, Google Search and Google Chrome are shown below.

- Very frequent **Facebook** users were significantly more likely to be female (50% compared to 35% male), aged 30–44 (54% compared to 39% 45–64 and 34% 65+), be C2 and DE (50% compared to 33% AB) and not have ad blockers (58% compared to 53%).
- Very frequent **Google Search** users were significantly more likely to be female (58% compared to 55% male), aged 18-29 (67% compared to 57% 45–64 and 43% 65+), be C1 (58% compared to 53% DE), be from a mixed or Asian background (75% compared to 55% White) and have a higher household income (62% above £40k compared to 54% below £40k).
- Very frequent **Google Chrome** users were significantly more likely to be aged 30-44 (53% compared to 34% 65+), be C1 (49% compared to 42% DE), be from an Asian background (62% compared to 44% White) and have a higher household income (51% above £60k compared to 44% £20k–£40k).

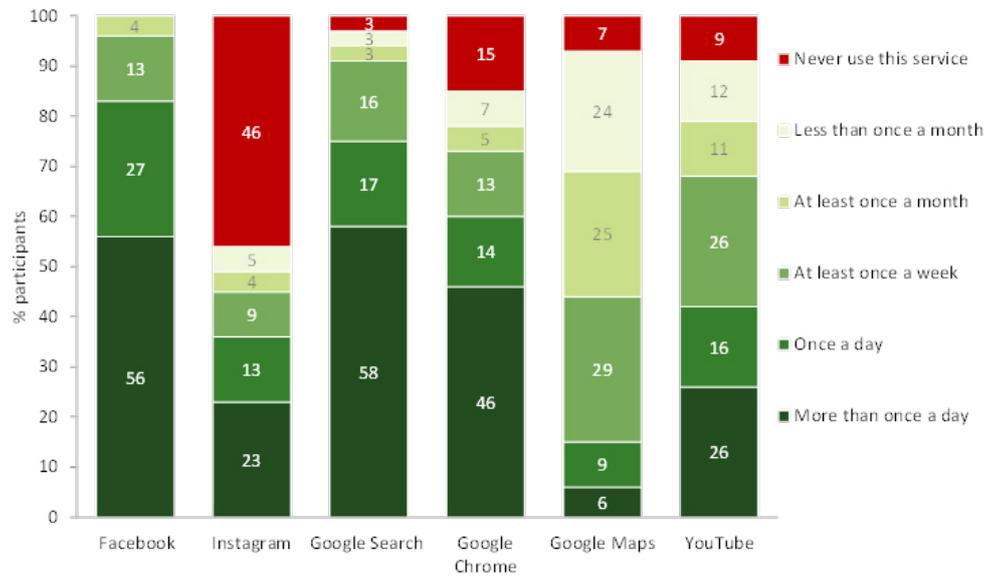
If both Google and Facebook<sup>17</sup> services were used there was a random allocation between the two.

For the Facebook sample 83% used Facebook at least once a day. Google Search was next most used (75% at least once a day) followed by Google Chrome (60%) and YouTube (42%). Instagram was used by fewest participants, with 46% not using the service within the preceding three months (see Figure 11). Our results for the use of Facebook services by Facebook users are comparable to the results obtained on the frequency of Facebook usage by the Which? quantitative survey discussed in Section 2.2. The Which? survey showed that 88% of Facebook users used Facebook at least once a day, followed by 10% who used it at least once a week and only 2% who used it at least once a month.

16 At the 95% confidence level

17 Unless just Instagram

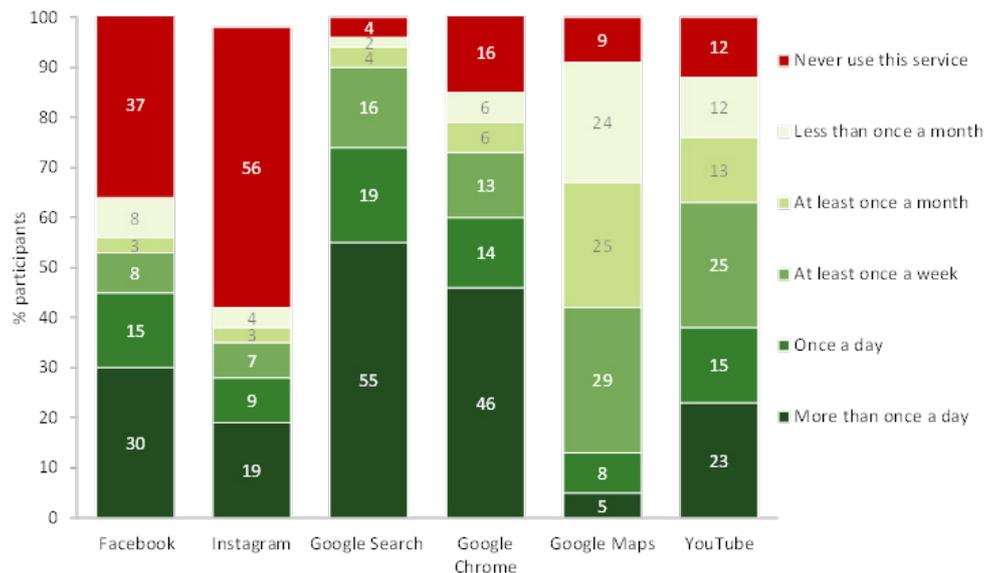
**Figure 11:** Use of Facebook and Google services in the preceding three months – Facebook sample



Base: Facebook sample 2,005

For the Google sample 74% used Google Search at least once a day. Google Chrome was next most used (60% at least once a day), followed by Facebook (45%) and YouTube (38%). Instagram was used by fewest participants, with 56% not using the service within the preceding three months. 37% of the Google sample did not use Facebook within the preceding three months. See Figure 12.

**Figure 12:** Use of Facebook and Google services in the preceding three months – Google sample

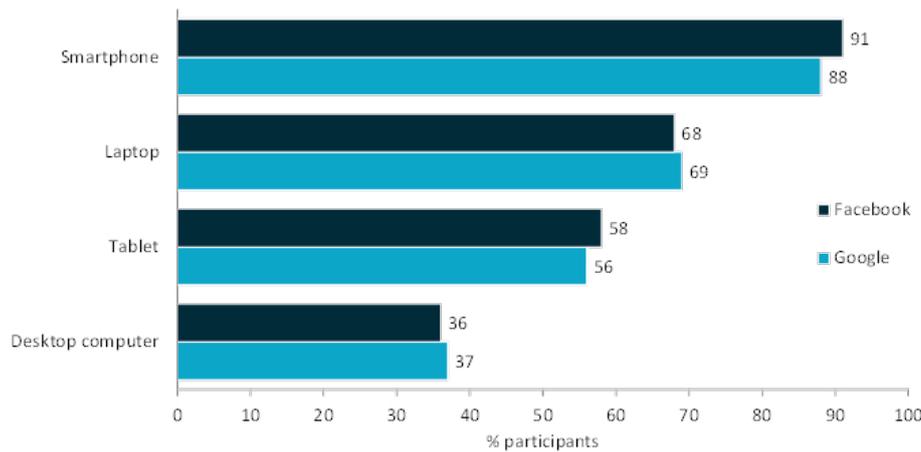


Base: Google sample 2,009

## Device usage

Overall, about nine tenths used a Smartphone (91%<sup>18</sup> Facebook sample, 88% Google sample). The next most used devices were a laptop (68% Facebook sample, 69% Google sample) and a tablet (58% Facebook sample, 56% Google sample). Desktop computers were least used with just over a third of each sample. See Figure 13.

**Figure 13:** Which of the following devices do you use?



Base: Facebook 2,005; Google 2,009

The main statistically significant differences<sup>19</sup> for the samples using Smartphones, Tablets, Laptops and Desktop computers are shown below.

- **Smartphone** users were significantly more likely to be female (91% compared to 87% male), younger (98% aged 18–29 and 96% aged 30–44 compared to 78% 65+) and have a higher household income (94% £40–60k and 93% over £60k compared to 85% less than £20k).
- **Tablet** users were significantly more likely to be older (65% aged over 65 compared to 52% 18–29), to be AB (62%) or C2 (58%) rather than DE (52%) and have a higher household income (64% over £60k, 61% £40–60k compared to 49% less than £20k).
- **Laptop** users were significantly more likely to be younger (76% aged 18–29 compared to 62% 65+), to be AB (76%) or C1 (72%) rather than C2 (63%) or DE (61%), have a higher household income (76% over £60k, 74% £40–60k compared to 62% less than £20k) and have ad blockers (73% compared to 67%).
- **Desktop computer** users were significantly more likely to be male (46% compared to 28% female), to be older (47% aged over 65 compared to 27% aged 18–29 and 32% aged 30–44), to be AB (45%) rather than DE (27%), have a higher household income (43% over £60k, 39% £40–60k, 38% £20–40k compared to 39% less than £20k) and have ad blockers (44% compared to 32%).

<sup>18</sup> This is significantly more than the 88% for the Google sample at the 95% confidence level

<sup>19</sup> At the 95% confidence level

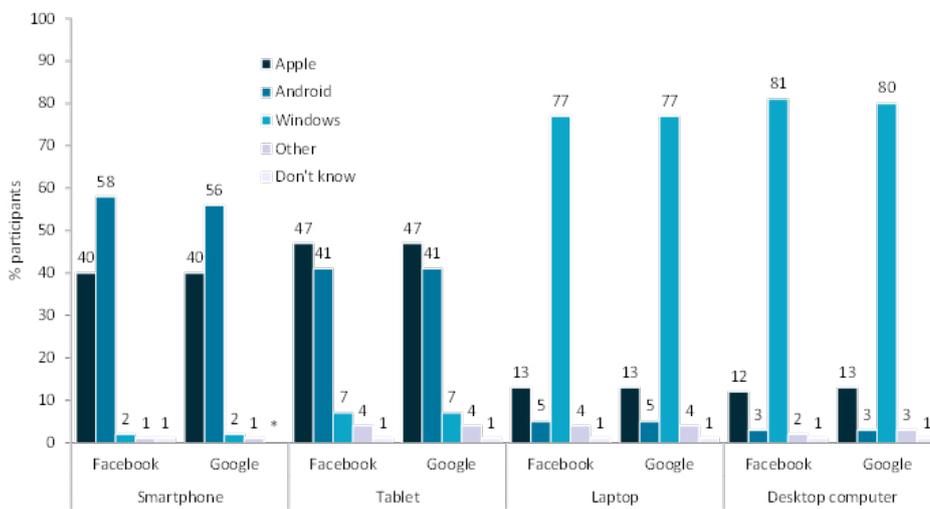
### Operating system

Participants were then asked which operating system they used on each of the devices that they used. If they used more than one on each device they were asked to think about the device they use most frequently for personal use.

There were very similar levels of operating system use by device between the two samples.

For Smartphones, 58% of the Facebook sample and 56% of the Google sample used Android and 40% of each sample used Apple. For laptops 77% of both samples used Windows and 13% of both samples used Apple. For tablets 47% of both samples used Apple and 41% of both samples used Android. There was a similar story for desktop computers as for laptops with about four fifths of both samples using Windows and 12–13% using Apple. See Figure 14.

**Figure 14:** Operating system used by device



Base: Facebook sample: Smartphone 1,819, Tablet 1,153, Laptop 1,354, Desktop computer 718;  
 Google sample: Smartphone 1,763, Tablet 1,130, Laptop 1,386, Desktop computer 738

Users of Apple on smartphones were significantly<sup>20</sup> more likely to have high incomes: 57% over £60k compared to 37% £20-40k and 31% under £20k. Users of Android on smartphones were significantly more likely to have low incomes: 65% under £20k and 61% £20-40k compared to 41% over £60k.

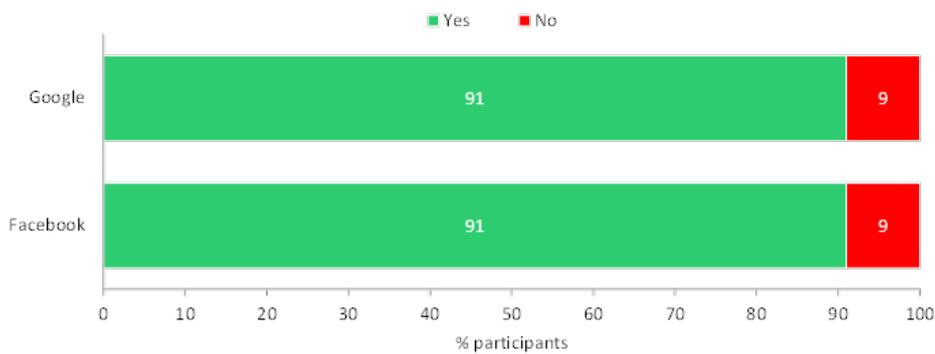
Interestingly, users of Apple on smartphones were significantly **less** likely to have ad blockers than to have them: 37% of Apple smartphone users had ad-blockers compared to 43% who did not. By contrast, users of Android on smartphones were significantly **more** likely to have ad blockers: 59% compared to 55%.

Users of Apple on tablets were significantly more likely to have high incomes: 56% over £60k and 57% £40-60k compared to 39% under £20k. They were also significantly **less** likely to have ad blockers 45% compared to 50%.

### 4.3 Awareness of and Attitudes towards Targeted Adverts

Awareness that data is collected, and used to target advertisements that are relevant, was high, with 91% claiming awareness for both samples (see Figure 15). Our results for the Facebook sample are comparable to the results of the Which? quantitative survey discussed in Section 2.2. The results of the Which? Survey showed that the vast majority (92%) of Facebook users in their sample reported to be aware of first-party data collection methods used by Facebook to inform targeted advertising. However, only two-thirds (66%) and a fifth (18%) of the respondents reported awareness of online tracking and customer lists.

**Figure 15:** Are you aware that Facebook/Google uses data that it collects about you to target adverts that it thinks are relevant to you?



Base: Facebook 2,005; Google 2,009

Younger participants were significantly<sup>21</sup> more likely to be aware than older participants: 95% 18–29 compared to 89% 65+. ABs (95%) and C1s (92%) were significantly more likely to be aware than C2s (88%) and DEs (87%). Those with higher household incomes were significantly more likely to be aware than those on lower incomes: 96% £60k or more compared to 88% less than £20k. Finally, those with ad blockers were significantly more likely to be aware than those without: 95% compared to 91%.

For laptop users, those with Apple operating system were significantly more likely than those with Windows operating system to be aware: 96% compared to 91%. There were no other significant differences by device and operating system.

#### Usefulness of targeted adverts

Half of the Facebook sample and 45% of the Google sample reported that they found targeted adverts useful (Facebook: 18%<sup>22</sup> very often 32% fairly often; Google: 14% very often, 31% fairly often).

For the Facebook sample 15% stated they never found targeted adverts useful. For the Google sample a similar proportion (16%) stated they never found targeted adverts useful. See Figure 16.

Users of Apple on Smartphones were significantly more likely to find targeted adverts useful very often than those on Android: 21% compared to 15%.

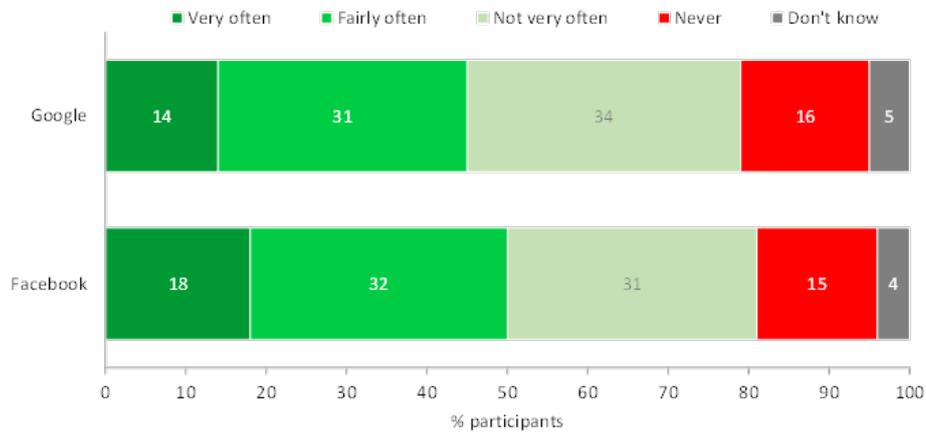
Users of Apple on laptops were significantly more likely to find targeted adverts useful very often than those on Windows: 25% compared to 14%.

<sup>21</sup> At the 95% confidence level

<sup>22</sup> This is significantly more than the 14% for the Google sample at the 95% confidence level

Similarly, users of Apple on desktop computers were significantly more likely to find targeted adverts useful very often than those on Windows: 22% compared to 13%.

**Figure 16:** How often do you find advertisements on (Facebook/Google) that are targeted (i.e. intended to be relevant) useful?



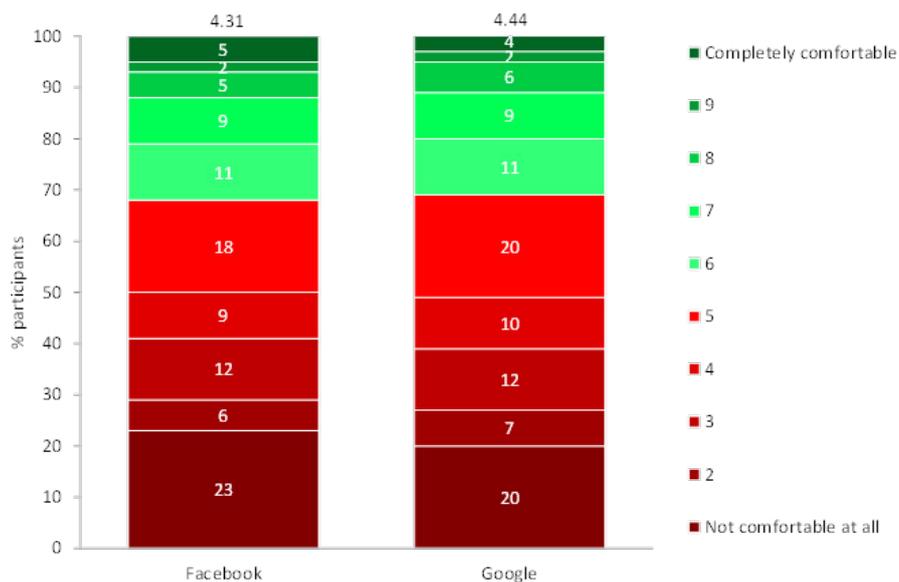
Base: Facebook 2,005; Google 2,009

### Comfort sharing data for personalised advertising

Participants were asked how comfortable they were with sharing their personal data with Facebook/Google for the purpose of targeted advertising.

There was a broad spread of comfort levels with data sharing with Facebook and Google. However, higher proportions reported lower levels of comfort than those who said they were comfortable with a mean comfort level of 4.31 for Facebook and 4.44 for Google (where 10 is completely comfortable and 1 not comfortable at all: 5.5 is the midpoint). See Figure 17.

**Figure 17:** In general, how comfortable are you sharing your personal data with (Facebook/Google) for the purpose of targeted advertising?



Base: Facebook 2,005; Google 2,009

The mean scores were significantly<sup>23</sup> higher (i.e. indicating they were more comfortable) for those without ad blockers than those with: 4.68 compared to 4.27.

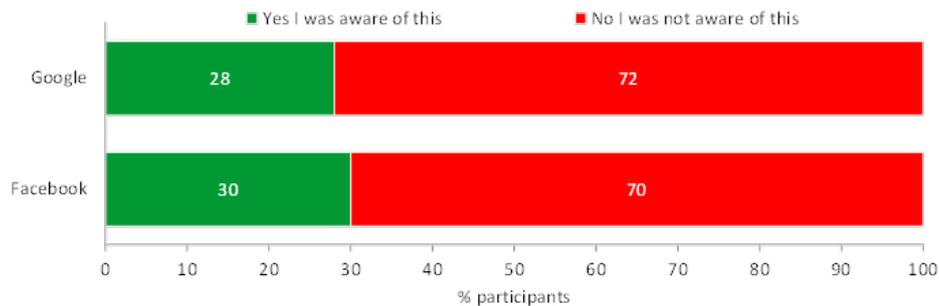
The mean scores were significantly higher for those on higher incomes: 4.90 for £60k or more compared to 4.28 for less than £20k and 4.34 for £20–40k.

The mean scores were significantly higher for younger participants: 5.33 for 18–29 and 5.02 for 30–44 compared to 3.65 for 65 or older.

**Awareness that Apple brought in new privacy changes**

Overall, there was low awareness that in April 2021 Apple brought in new privacy changes which gives users the choice of whether the apps on their phone can track their activity across other company’s apps and websites: 28% of the Google sample and 30% of the Facebook sample were aware.

**Figure 18:** Were you aware that in April 2021 Apple brought in new privacy changes which gives users the choice of whether the apps on their phone can track their activity across other company’s apps and websites?



Base: Facebook 2,005; Google 2,009

As would be expected users of Apple operating systems were much more aware of the new privacy changes than users of other operating systems. All the differences shown below were statistically significant.

**Table 3:** Awareness of the new Apple privacy changes by main operating systems by device

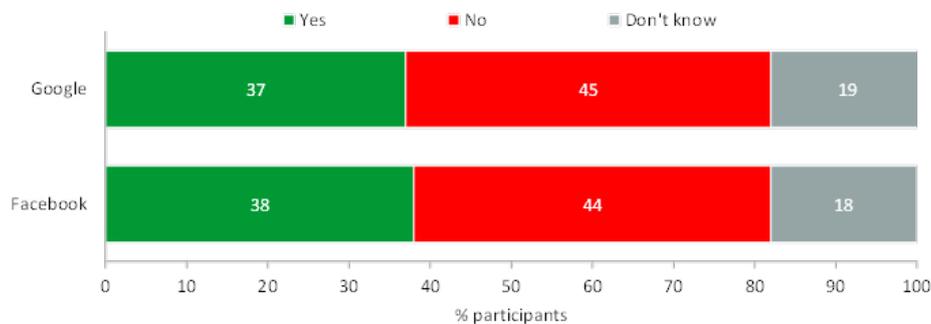
	Smartphone		Tablet		Laptop		Desktop	
	Apple	Android	Apple	Android	Apple	Windows	Apple	Windows
Aware	42%	23%	40%	27%	51%	27%	49%	26%
Not aware	58%	77%	60%	73%	49%	73%	51%	74%
<b>Base</b>	<b>1438</b>	<b>2041</b>	<b>1075</b>	<b>936</b>	<b>358</b>	<b>2109</b>	<b>183</b>	<b>1177</b>

23 At the 95% confidence level

## 4.4 Privacy settings

Over a third of each sample had ad blockers set up on one or more of their devices: 38% Facebook and 37% Google. Just under a fifth of both samples did not know. See Figure 19.

**Figure 19:** Thinking about the devices that you use do you have ad blockers set up on any of these devices?



Base: Facebook 2,005; Google 2,009

Younger participants were significantly<sup>24</sup> more likely to have ad blockers than older participants: 43% 18–29 compared to 37% 45–64 and 36% 65+. Males were significantly more likely to have ad blockers than females: 41% compared to 35%. ABs (42%) were significantly more likely to have ad blockers than DEs (34%). Those with higher household incomes were significantly more likely to have ad blockers than those on lower incomes: 46% £60k or more compared to 35–38% less than £60k.

### Awareness of data preference settings

Two-fifths of both samples had no awareness that they could change their data preference settings. For the Facebook sample 22% said they were aware and had changed their settings,<sup>25</sup> 13% had reviewed but had not changed their settings and 21% were aware but had not gone in to look at their settings. Our results for the Facebook sample are comparable to the results of the Which? quantitative survey discussed in Section 2.2. The results of the Which? survey showed that only 11% of Facebook users in their sample were aware of the ‘off-Facebook activity’ setting that they could use to control and monitor online tracking.

For the Google sample 17% said they were aware and had changed their settings, 14% had reviewed but had not changed their settings and 26% were aware but had not gone in to look at their settings.

Users of Apple on laptops and desktop computers were significantly more likely to have changed their settings than users of Windows:

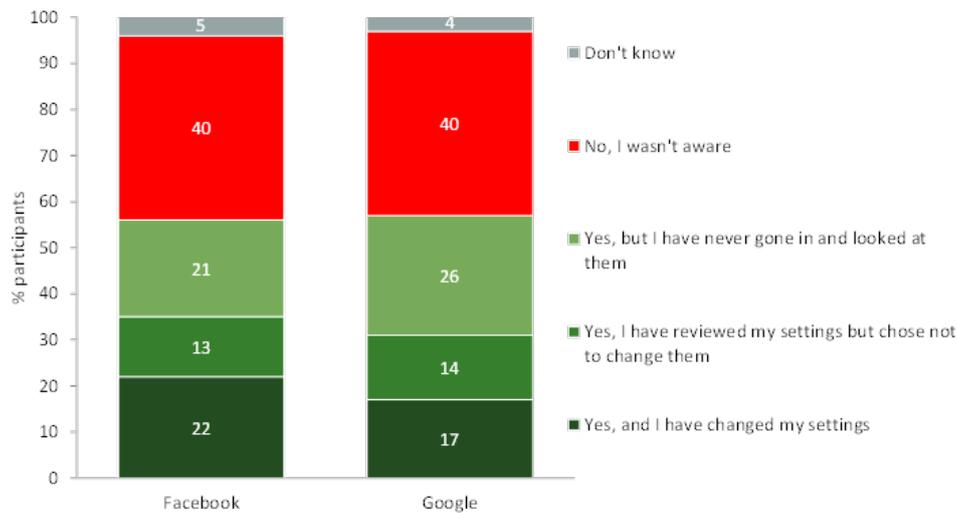
- Tablets: 27% compared to 19%
- Laptops: 30% compared to 20%

Users of Windows on tablets were significantly more likely to be not aware that they could change their settings than Apple users: 39% compared to 30%. There were no other significant differences by device and operating system.

<sup>24</sup> At the 95% confidence level

<sup>25</sup> This is significantly more than the 17% for the Google sample at the 95% confidence level.

**Figure 20:** Were you aware that you could change your settings relating to your data and how (Facebook/Google) can use that information to inform the adverts that you see when you are using (Facebook/Google)?



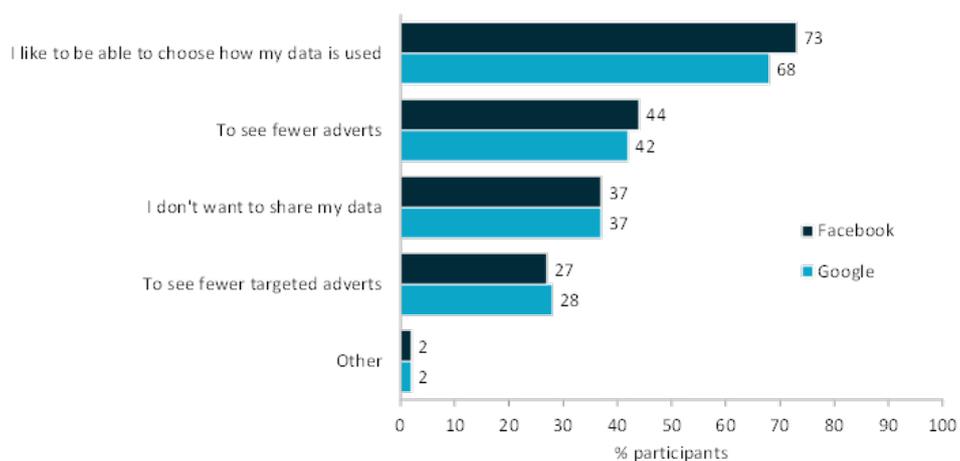
Base: Facebook 2,005; Google 2,009

Nearly a third (32%) who had ad blockers installed had changed their settings compared to 12% who did not have ad blockers.

Nearly a half (46%) who do not have ad blockers were not aware they could change their settings compared to 27% who have ad blockers.

The 434 (22%) of the Facebook sample and 335 (17%) of the Google sample who said that they had changed their settings were asked why. The main reason given by both samples was they liked to be able to choose how their data was used (73% Facebook, 68% Google). Also important was to see fewer adverts (44% Facebook, 42% Google) and not wanting to share their data (37% each). It should be noted that the change in settings would not lead to fewer adverts – this is a misunderstanding by some users. See Figure 21.

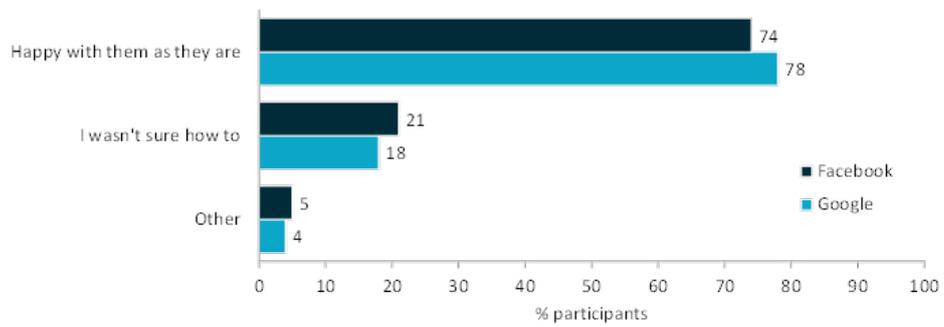
**Figure 21:** Why did you change your settings?



Base: Those who changed settings: Facebook sample 434, Google sample 335

The 261 (13%) of the Facebook sample and 277 (14%) of the Google sample who said that they had reviewed their settings but did not change them were asked why they did not change their settings. The main reason given by both samples was that they were happy with them as they were (74% Facebook, 78% Google). About a fifth (21% Facebook, 18% Google) said they were not sure how to change their settings.

**Figure 22:** Why did you choose not to change your settings?



Base: Those who reviewed settings but did not change them: Facebook sample 261, Google sample 277

## 4.5 Informed Attitudes towards Data Sharing

Participants were then informed about three of the ways that Facebook/Google collects information and asked them how they felt about each of them.

The three ways were:

- First party data collection
- Third party online tracking
- Third party customer lists

For each, participants for each sample group were shown a screen with information about these aspects:

**Figure 23:** Screenshots showing information about data collection: Facebook

**First-party data collection**

One way Facebook collects information about you is by observing what you do on Facebook and Instagram (which Facebook also owns) and from your profile information. If you have accounts with both, information about what you do on Instagram will influence what you see on Facebook. Facebook then uses this data to try to make the adverts you see at Facebook more relevant.

It collects information on things like:

- what posts and pages you “like”
- information you have added to your Facebook and Instagram profile
- places that you “check into”
- ads and other content that you watch or engage with (e.g. click on, videos that you watch).

By collecting this information Facebook knows things like your age, gender, location and what your interests are. This is then used to target you with adverts that Facebook thinks you will be interested in.

**Third-party: online tracking**

Facebook also collects data about you from other organisations i.e. other businesses or charities that have a relationship with Facebook. Facebook uses this data to try to make the adverts you see on Facebook more relevant.

One of the ways they collect information from other organisations is to *track your activity* when you’re on websites and apps other than Facebook. They do this using tools like cookies, or if you log-in using your Facebook log-in.

Information is sent back to Facebook about web pages you’ve looked at, whether you’ve purchased something or put something in your online basket.

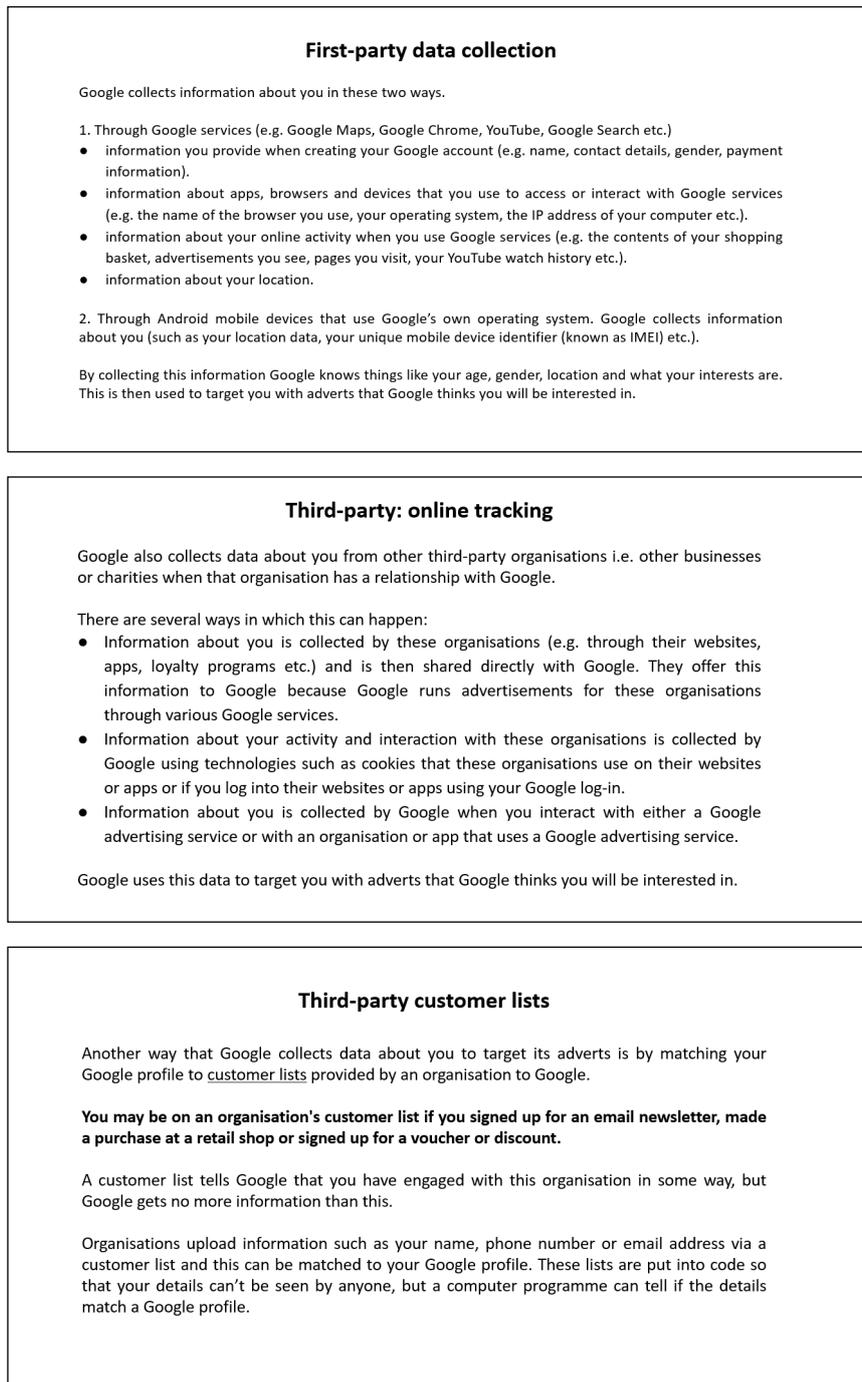
**Third-party customer lists**

Another way that Facebook collects data about you to target its adverts is by matching your Facebook profile to customer lists provided by an organisation to Facebook.

**You may be on an organisation’s customer list if you signed up for an email newsletter, made a purchase at a retail shop or signed up for a voucher or discount.**

A customer list tells Facebook that you have engaged with this organisation in some way, but Facebook gets no more information than this.

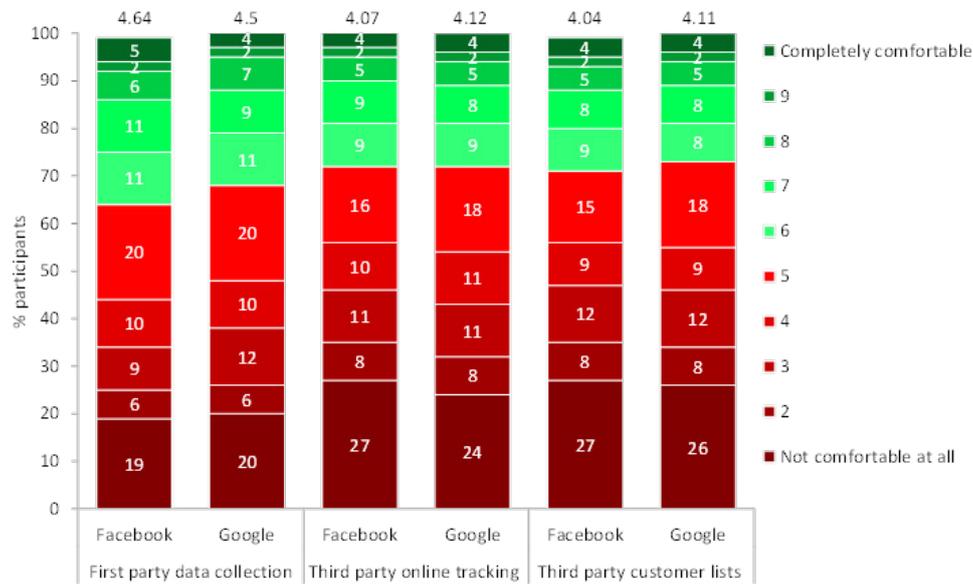
Organisations upload information such as your phone number or email address via a customer list and this can be matched to your Facebook profile. These lists are put into code so that your details can’t be seen by anyone, but a computer programme can tell if the details match a Facebook profile.

**Figure 24:** Screenshots showing information about data collection: Google

For both samples comfort levels with sharing data were slightly lower for 'Third party online tracking' and 'Third party customer lists' than for 'First party data collection'. This result is in line with the results obtained by the Which? quantitative survey (discussed in Section 2.2) which showed that first-party data collection methods were more likely to be accepted by Facebook users than third-party methods.

Higher proportions reported lower levels of comfort than those who said they were comfortable for all three aspects with mean comfort levels of 4.04–4.64 for the Facebook sample and 4.11–4.64 for the Google sample (where 10 is completely comfortable and 1 not comfortable at all: 5.5 is the midpoint). See Figure 25.

**Figure 25:** How comfortable are you for Facebook/Google to collect first party data, third party online tracking and third party customer lists in order to inform targeted adverts?



Base: Facebook 2,005; Google 2,009

Those without ad blockers gave significantly<sup>26</sup> higher mean scores (indicating higher comfort levels) for:

- First party data collection: 4.89 compared to 4.44
- Third party online tracking: 4.41 compared to 3.94
- Third party customer lists: 4.39 compared to 3.96.

Significantly higher mean scores were also given for all three aspects by those on higher incomes, by those of Black and Mixed ethnicities compared to White and by under 44 year olds compared to over 45 year olds.

Those who were aware that in April 2021 Apple brought in new privacy changes were more comfortable than those who were not aware:

- First party data collection: 5.41 compared to 4.24.
- Third party online tracking: 5.03 compared to 3.72.
- Third party customer lists: 5.03 compared to 3.69.

Those who had reviewed their privacy settings but chose not to change them were more comfortable than those who were aware and those not aware:

- First party data collection: 6.01 compared to 3.02 aware and changed, 4.79 aware but not looked at settings and 4.07 not aware.
- Third party online tracking: 5.52 compared to 3.95 aware and changed, 4.23 aware but not looked at settings and 3.61 not aware.
- Third party customer lists: 5.48 compared to 3.96 aware and changed, 4.17 aware but not looked at settings and 3.62 not aware.

For smartphone users, significantly higher mean scores were given by Apple than Android users for Third party customer lists: 4.33 compared to 3.98. There were no other significant differences by device and operating system.

26 At the 95% confidence level

# 5. Contingent Valuation Analysis and Findings

## 5.1 Introduction

This section presents findings on the choices made by participants in the contingent valuation exercises. All the results are segmented by uninformed vs informed scenario, service (Facebook/Google), and payment vehicle type (fee/reward). The section begins by presenting findings from a descriptive analysis of choices (Section 5.2). The main results on values for the CRR are then presented in Section 5.3. These results are themselves derived from an econometric analysis of the choice data. Details of the econometrics, including model specifications, testing procedures and estimation results, are presented in Appendices C and D. Section 5.4 presents an analysis of choices between data sharing alternatives.

## 5.2 Descriptive Analysis

### Uninformed and Informed choices without reward/fee

Table 4 shows the proportions of participants in the full sample who chose generic versus targeted adverts under the uninformed scenario and the informed scenario when there was no associated reward or fee. Under the uninformed scenario, 57% chose generic adverts and the remaining 43% chose targeted adverts. However, once they were informed of how Google and Facebook collect personal data from them, the proportion choosing generic adverts increased to 73% and the proportion choosing targeted adverts decreased to 27%. The difference in proportions between uninformed and informed choices is statistically significant at the 1% level. This result is in line with the findings of the quantitative study conducted by Which? discussed in Section 2.2. The survey found that once participants were informed about the data collection methods used to inform targeted adverts, there was a shift from a slim majority preferring targeted adverts to a plurality preferring generic adverts.

**Table 4:** Uninformed vs. informed choices without payment (full sample)

Uninformed choice	Informed choice		
	Generic	Targeted	Total
Generic	51%	7%	<b>57%</b>
Targeted	22%	20%	<b>43%</b>
Total	<b>73%</b>	<b>27%</b>	<b>100%</b>

Base: 4,014. Percentages may not add up to 100% due to rounding.

The proportions choosing ‘Generic’ do not differ statistically between user groups (Facebook/Google) in both uninformed and informed choices (two-sample test of proportions,  $p$ -values > 0.29). See Table 5.

**Table 5:** Uninformed vs. informed choices without payment (by service)

		Informed choice		
		Generic	Targeted	Total
Facebook (Base: 2,005)	Uninformed choice			
	Generic	51%	7%	<b>58%</b>
	Targeted	22%	20%	<b>42%</b>
<b>Total</b>		<b>73%</b>	<b>27%</b>	<b>100%</b>
Google (Base: 2,009)	Generic	51%	6%	<b>57%</b>
	Targeted	22%	20%	<b>43%</b>
	<b>Total</b>	<b>74%</b>	<b>26%</b>	<b>100%</b>

Percentages may not add up to 100% due to rounding.

**Impact of fee/reward on Uninformed choices**

Table 6 shows the proportion of uninformed choices without any fee/reward against similarly uninformed choices but with an associated reward or fee. It shows that 43% chose targeted adverts without any associated financial incentive; however, this increased to 84% once a fee/reward was offered. This indicates that the financial incentive worked as expected to encourage around 40% of participants to switch from generic to targeted adverts. Moreover, only a small proportion (3%) made the counter-intuitive opposite switch from ‘Targeted’ to ‘Generic’. The difference in the proportion choosing ‘Generic’ without and with payment is very strongly statistically significant in the full sample and for each of the two payment vehicles (fee/reward) and services (Facebook/Google).

**Table 6:** Uninformed choices with and without payment (full sample)

Without fee/reward	With fee/reward		
	Generic	Targeted	Total
Generic	13%	44%	<b>57%</b>
Targeted	3%	40%	<b>43%</b>
<b>Total</b>	<b>16%</b>	<b>84%</b>	<b>100%</b>

Base: 4,014; based on responses to 1<sup>st</sup> fee/reward question only, i.e., not including follow-up responses. Percentages may not add up to 100% due to rounding.

The difference in the proportions choosing ‘Targeted’ across fee/reward treatments before being shown the question including a reward or fee is small, and it is not statistically significant (see Table 7). Consistent with expectations, fees have a stronger impact than rewards (47% switch with the fee compared to 35% with the reward), and the proportion choosing ‘Targeted’ is substantially higher with the fee than with the reward.

**Table 7:** Uninformed choices with and without payment (by payment vehicle)

	Without fee/reward	With fee/reward		
		Generic	Targeted	Total
<b>Fee</b> (Base: 2,000)	Generic	6%	50%	<b>57%</b>
	Targeted	3%	41%	<b>44%</b>
	<b>Total</b>	<b>9%</b>	<b>91%</b>	<b>100%</b>
<b>Reward</b> (Base: 2,014)	Generic	20%	38%	<b>58%</b>
	Targeted	3%	39%	<b>42%</b>
	<b>Total</b>	<b>23%</b>	<b>77%</b>	<b>100%</b>

Note: Based on responses to 1<sup>st</sup> fee/reward question only, i.e., not including follow-up responses.  
Percentages may not add up to 100% due to rounding.

Table 8 shows the proportions of choices segmented by service (Facebook/Google). There are again no statistically significant differences across services, both without and with fee/reward. Both groups of users showed a substantially increased proportion choosing targeted adverts when a financial incentive was offered, as expected.

**Table 8:** Uninformed choices with and without payment (by service)

	Without fee/reward	With fee/reward		
		Generic	Targeted	Total
<b>Facebook</b> (Base: 2,005)	Generic	14%	44%	<b>58%</b>
	Targeted	3%	40%	<b>42%</b>
	<b>Total</b>	<b>17%</b>	<b>83%</b>	<b>100%</b>
<b>Google</b> (Base: 2,009)	Generic	12%	45%	<b>57%</b>
	Targeted	3%	40%	<b>43%</b>
	<b>Total</b>	<b>15%</b>	<b>85%</b>	<b>100%</b>

Note: Based on responses to 1<sup>st</sup> fee/reward question only, i.e., not including follow-up responses.  
Percentages may not add up to 100% due to rounding.

Overall, in Table 6 to Table 8, we find, as expected, that the proportion of participants choosing targeted adverts increases after they are presented with a financial incentive.

### Impact of fee/reward on Informed choices

Table 9 shows the proportion of informed choices between generic and targeted adverts without any reward/fee against those same choices with an associated reward/fee. The proportion choosing targeted adverts increased from 27% without any associated reward/fee to 68% when associated with a reward/fee. As under uninformed choices, around 40% switched from ‘Generic’ to ‘Targeted’ when given an incentive for switching.

The proportion choosing ‘Generic’ was substantially higher under informed than under uninformed choices both without (73% vs. 57%) and with payment (32% vs. 16%), and these differences are strongly statistically significant. Information about how Google/Facebook collect data thus led to a greater proportion choosing generic adverts despite the financial incentive for switching.

**Table 9:** Informed choices with and without fee/reward (full sample)

Without fee/reward	With fee/reward		
	Generic	Targeted	Total
Generic	30%	43%	<b>73%</b>
Targeted	2%	24%	<b>27%</b>
<b>Total</b>	<b>32%</b>	<b>68%</b>	<b>100%</b>

Base: 4,014; based on responses to 1<sup>st</sup> fee/reward question only, i.e., not including follow-up responses. Percentages may not add up to 100% due to rounding.

Table 10 segments the proportions by whether a reward or a fee was offered as a payment vehicle for incentivising choice of targeted adverts. Unlike for uninformed choices, the proportions choosing ‘Generic’ differ statistically at the 1% level between fee and reward treatments even in the question before the fee/reward is offered, but the difference in proportions is not very large.<sup>27</sup>

Fees have an even stronger impact than rewards compared to uninformed choices in terms of participants switching from ‘Generic’ to ‘Targeted’, with 59% switching from ‘Generic’ to ‘Targeted’ when asked to pay a fee to avoid receiving targeted adverts compared to only 28% among those who were offered a reward for choosing ‘Targeted’.

**Table 10:** Informed choices with and without payment (by payment vehicle)

	Without fee/reward	With fee/reward		
		Generic	Targeted	Total
<b>Fee (Base: 2,000)</b>	Generic	17%	59%	<b>76%</b>
	Targeted	3%	22%	<b>24%</b>
	<b>Total</b>	<b>19%</b>	<b>81%</b>	<b>100%</b>
<b>Reward (Base: 2,014)</b>	Generic	43%	28%	<b>71%</b>
	Targeted	2%	27%	<b>29%</b>
	<b>Total</b>	<b>45%</b>	<b>55%</b>	<b>100%</b>

Note: Based on responses to 1<sup>st</sup> fee/reward question only, i.e., not including follow-up responses. Percentages may not add up to 100% due to rounding.

Again, as for uninformed choices, there were no statistically significant differences between Facebook and Google users. (See Table 11.)

27 Because of large sample size, which yields more precise estimates, even relatively small differences are flagged as statistically significant.

**Table 11:** Informed choices with and without payment (by service)

	Without fee/reward	With fee/reward		
		Generic	Targeted	Total
<b>Facebook</b> (Base: 2,005)	Generic	30%	43%	<b>73%</b>
	Targeted	2%	25%	<b>27%</b>
	<b>Total</b>	<b>32%</b>	<b>68%</b>	<b>100%</b>
<b>Google</b> (Base: 2,009)	Generic	30%	43%	<b>74%</b>
	Targeted	2%	24%	<b>26%</b>
	<b>Total</b>	<b>32%</b>	<b>68%</b>	<b>100%</b>

Note: Based on responses to 1<sup>st</sup> fee/reward question only, i.e., not including follow-up responses.

Percentages may not add up to 100% due to rounding.

### Impact of fee/reward size on choices

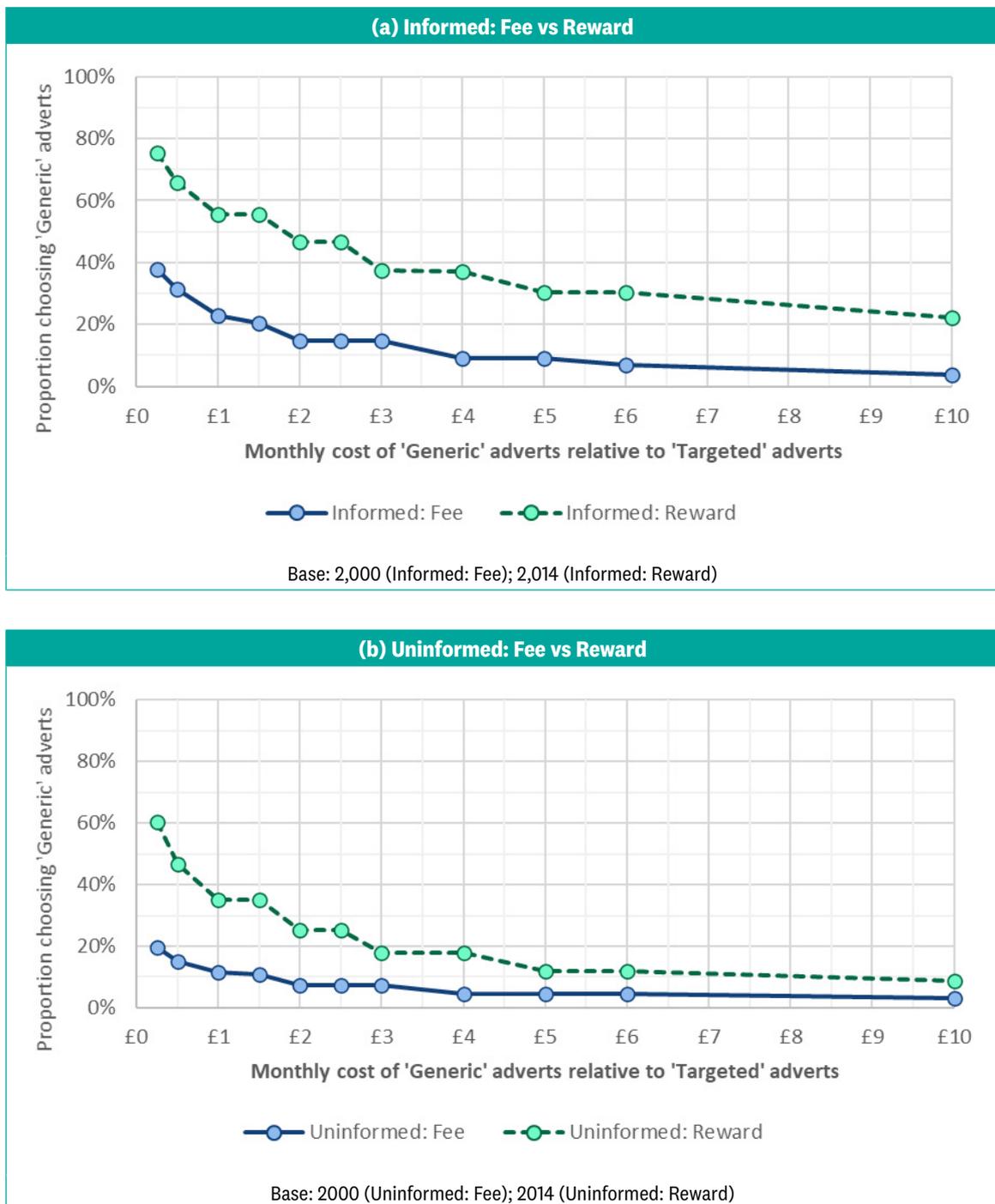
The following figures show the proportions choosing ‘Generic’ adverts by the cost of that option relative to the cost of ‘Targeted’ adverts, i.e., the amount of the reward or fee depending on which payment vehicle treatment a participant was randomly assigned to. The proportions shown are based on responses to the second and third contingent valuation questions of each set of uninformed/informed choices, in which participants would forego a reward/pay a fee if they chose not to share their data.

The downward sloping curves indicate that participants were more likely to choose generic adverts when cheaper than when more expensive. The proportion choosing ‘Generic’ at the highest cost amount of £10 per month was around 20% or less (depending on payment vehicle and information).

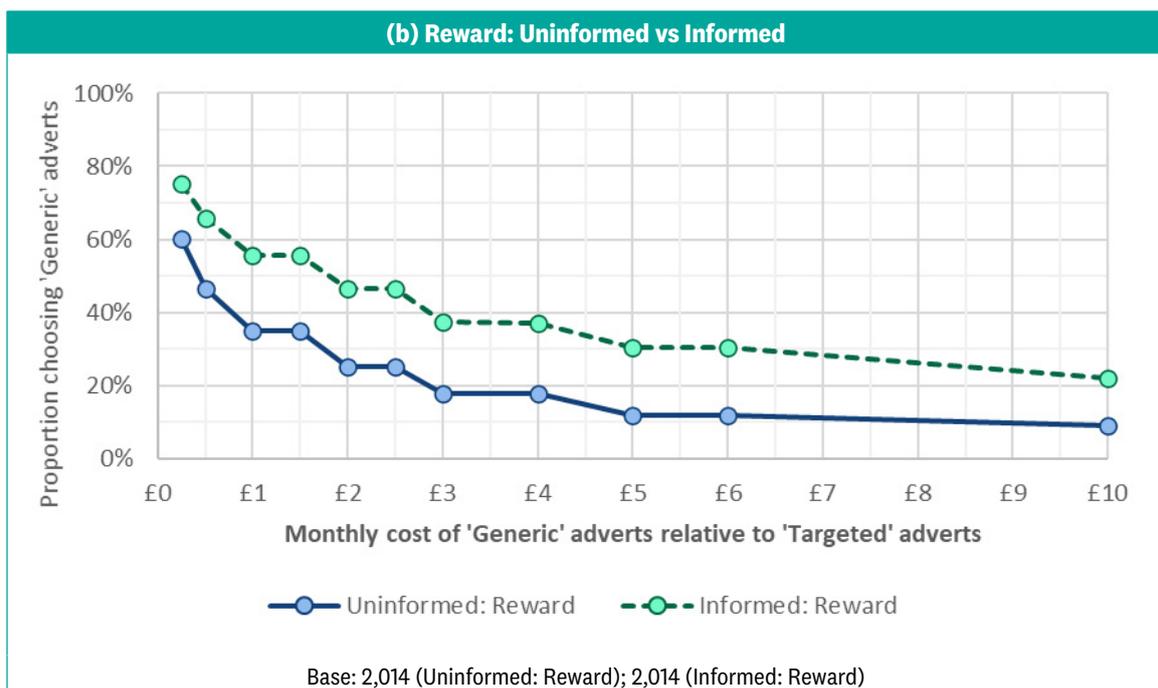
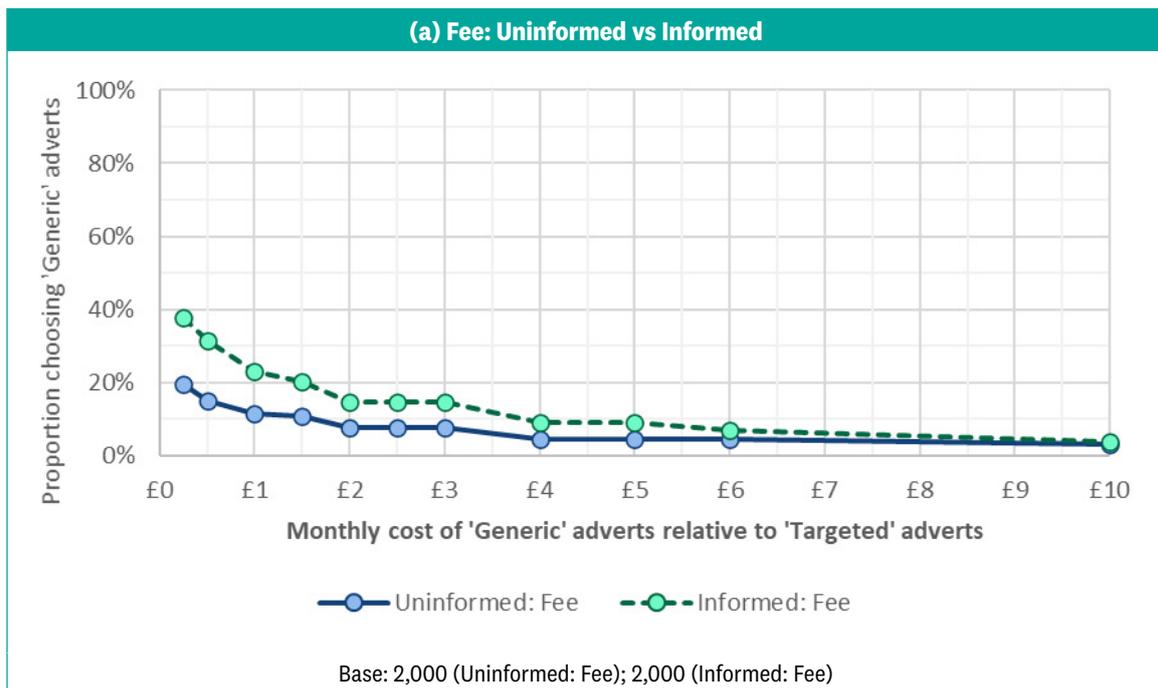
The shape and location of the curves corroborate our findings from the analysis of the proportions choosing ‘Generic’ vs ‘Targeted’ in the first and second contingent valuation questions of each set of uninformed/informed choices under different scenarios (see Table 4 to Table 11):

- At any given relative cost of the ‘Generic’ option, and especially at low to medium levels (£0.25 to £4 per month), participants were far more likely to choose ‘Generic’ when offered a reward for receiving ‘Targeted’ adverts than when asked to pay a fee for receiving ‘Generic’ adverts.
- A higher proportion of participants chose to see ‘Generic’ adverts under the informed scenario vis-à-vis the uninformed scenario across all cost amounts and subsamples.
- The proportions choosing ‘Generic’ in uninformed vs informed choices were relatively similar in the fee scenario compared to the reward scenario, i.e., information had a greater impact when participants were offered a reward for choosing ‘Targeted’.
- The proportions choosing ‘Generic’ under any scenario were very similar in the Facebook and Google user groups.

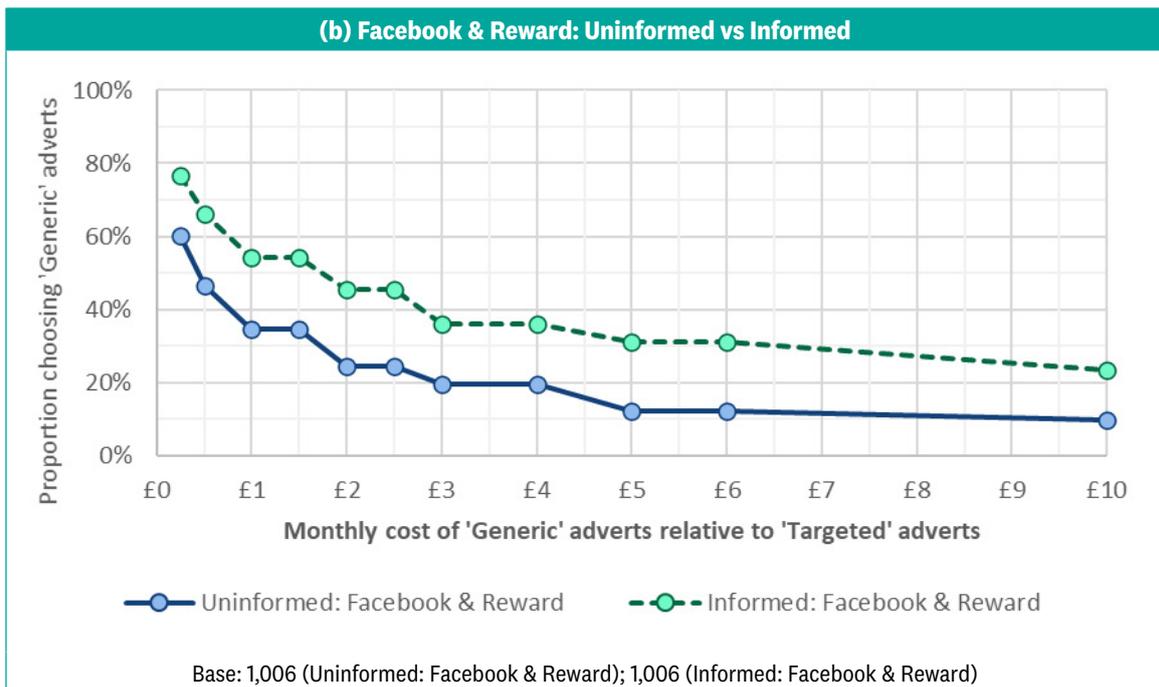
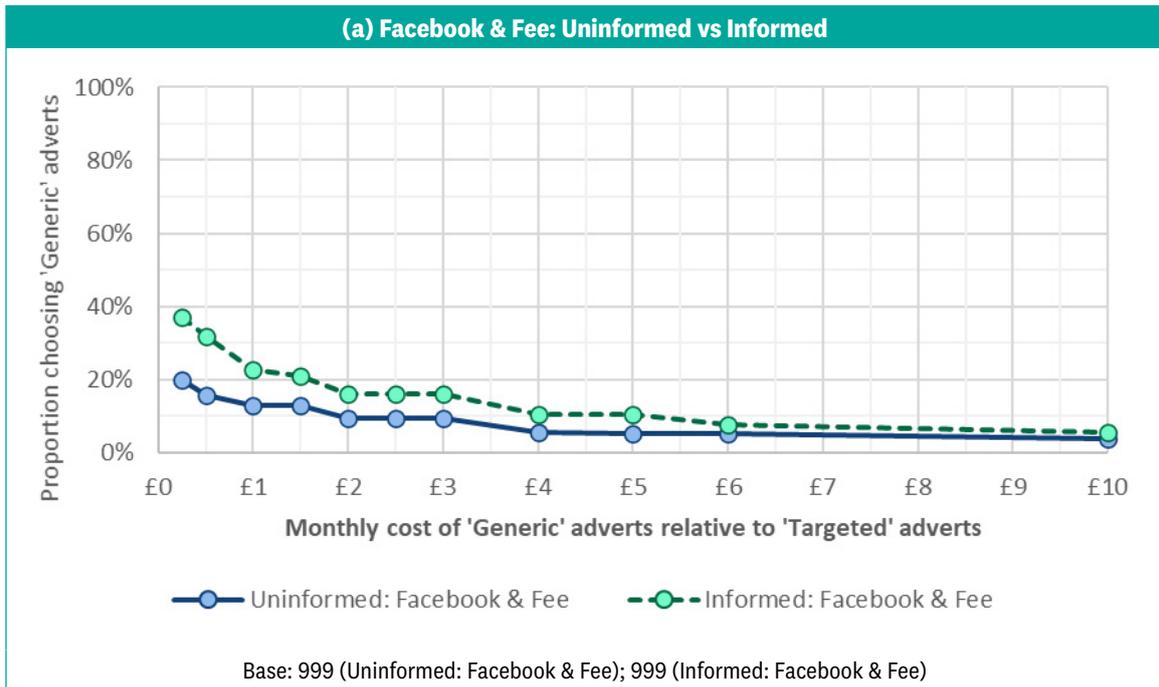
**Figure 26:** Estimated proportion choosing 'Generic' adverts by cost of 'Generic' adverts: Facebook and Google users



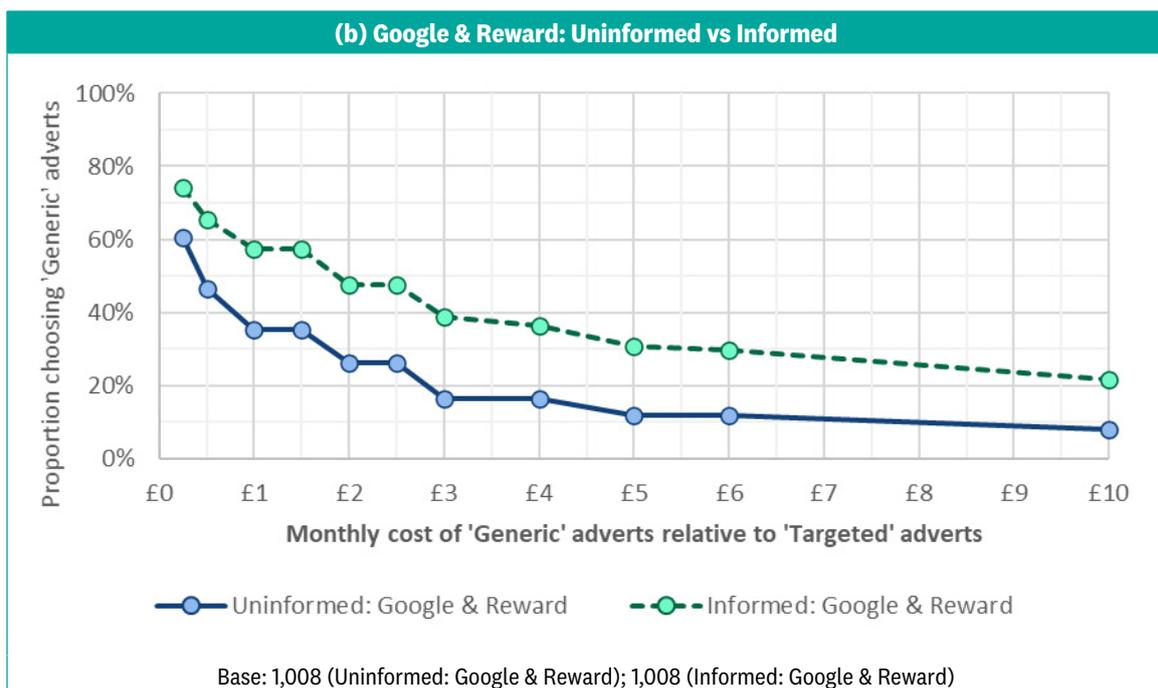
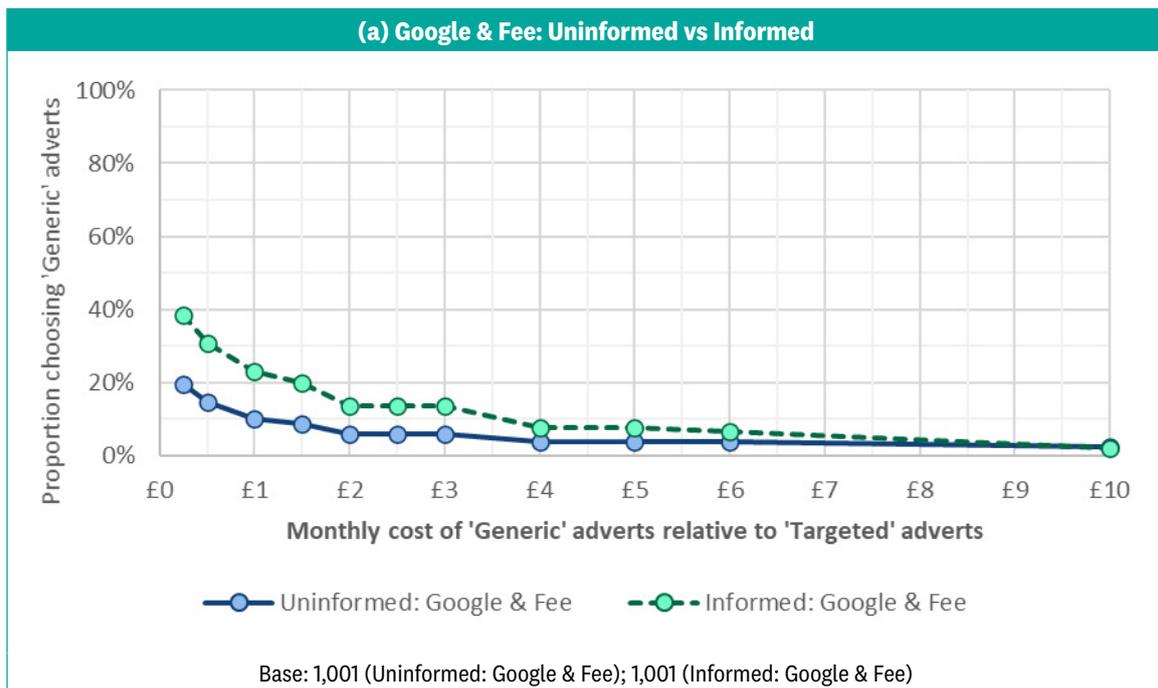
**Figure 27:** Estimated proportion choosing 'Generic' adverts by cost of 'Generic' adverts: Facebook and Google users



**Figure 28:** Estimated proportion choosing 'Generic' adverts vs cost of 'Generic' adverts: Facebook users



**Figure 29:** Estimated proportion choosing 'Generic' adverts vs cost of 'Generic' adverts: Google users



## 5.3 Valuation of the Choice Requirement Remedy

### Main estimates

As discussed in section 3.2, our approach to valuing the CRR is based on the premise that the CRR will expand the ‘choice set’ of Facebook and Google users by effectively giving them the choice not to share their data for targeted advertising. Our estimate is given by the mean change in expected consumer surplus arising from a change in the choice set. The mean value can be aggregated to obtain a total CRR value estimate for the whole of the UK.

We calculate the change in expected consumer surplus based on an econometric modelling analysis of the choice data. Appendix C describes the underlying choice model and our approach to valuing the CRR and presents our baseline models, sensitivity tests, and final models. Both our baseline and final models include all contingent valuation choices between ‘Generic’ and ‘Targeted’ adverts made by each participant and allow for heterogeneous preferences among participants regarding sharing of personal data for targeted advertising.

There is a degree of judgement in deciding which valuation scenario is most relevant to the appraisal of the CRR. Moreover, approaches and views vary amongst researchers regarding sample exclusions. Therefore, we explore the sensitivity of valuations of the CRR to a range of scenarios and assumptions. Table 12 shows our findings. We do not differentiate between Facebook and Google users as our segmentation analysis (see next section) showed no statistically significant differences between the two groups. This is in line with the findings in Section 4.3 which show that, overall, there was a similar spread of comfort levels with data sharing among both Facebook and Google users (see Figure 17).

**Table 12:** Mean value of the CRR: Baseline estimates, sensitivity tests, and final estimates (£/user/month)

Approach	Fee scenario		Reward scenario	
	Uninformed choices	Informed choices	Uninformed choices	Informed choices
(1) Baseline estimates (full sample)	£1.12 [£0.92 -£1.31]	£1.86 [£1.62 -£2.10]	£2.26 [£1.94 -£2.58]	£5.77 [£4.87 -£6.67]
(2) CRR valuation excluding fastest 10% of participants in terms of questionnaire completion time	£1.01 [£0.83 -£1.20]	£1.82 [£1.58 -£2.05]	£2.27 [£1.93 -£2.61]	£5.74 [£4.83 -£6.65]
(3) CRR valuation of those who were 'Very sure' or 'Somewhat sure' that they would make the same choices in a real choice situation	£1.19 [£0.96 -£1.43]	£1.95 [£1.66 -£2.24]	£2.36 [£2.01 -£2.71]	£5.66 [£4.72 -£6.61]
(4) CRR valuation of those who: - were not among the fastest 10% of participants; and - were 'Very sure' or 'Somewhat sure' of their choices; and - had not reviewed their data settings <sup>(2)</sup>	£0.72 [£0.60 -£0.85]	£1.56 [£1.36 -£1.75]	£2.38 [£1.95 -£2.81]	£5.68 [£4.61 -£6.75]
(5) <b>Weighted average of:</b> - CRR valuation of 'non-reviewers' <sup>(2)</sup> derived from (4) - Zero for those who had reviewed their data settings <sup>(3)</sup>	<b>£0.50</b> <b>[£0.41</b> <b>-£0.59]</b>	<b>£1.09</b> <b>[£0.95</b> <b>-£1.23]</b>	<b>£1.68</b> <b>[£1.38</b> <b>-£1.99]</b>	<b>£4.03</b> <b>[£3.26-</b> <b>£4.79]</b>

(1) All estimates refer to a CRR providing a dichotomous choice between 'Full data sharing' (and receiving targeted adverts) and 'No data sharing' (and receiving 'Generic' adverts). All values are in £/user/month. 95% Krinsky-Robb confidence intervals based on 1,000 parameter draws are shown in parentheses (see details in Appendix C). The underlying econometric models are shown in Appendix C, Table 24 and Table 26.

(2) 'Non-reviewers' include those who reviewed their data settings but were not sure how to change them.

(3) Excluding those who were not sure how to change their data settings.

Two key factors affecting valuations are the payment vehicle – fee or reward – used in the valuation scenario and the level of information about data collection by online platforms. Valuations in the reward scenario are between 2.0 and 3.7 times higher than in the fee scenario, while valuations based on informed choices are between 1.6 and 2.6 times higher than uninformed choices.

The baseline estimates are not very sensitive to excluding participants whose questionnaire completion time was substantially below average. The CRR valuations of those who were 'Very sure' or 'Somewhat sure' that they would make the same choices in a real choice situation were very similar to our baseline estimates. Allowing for differences in valuations between those who had reviewed/not reviewed their data settings leads to lower estimates of the value of the CRR for the latter group, but only in the 'Fee' treatment.

Our final estimates of user valuations of the CMA's proposed CRR which combine all the above changes to our baseline approach and, additionally, impute a zero valuation for those who have

changed their privacy settings or who, having reviewed their settings, decided not to change them,<sup>28</sup> are 25% to 55% lower than our baseline estimates, and range between £0.50 and £4.03 per user per month depending on whether the valuation is based on an informed or an uninformed question, whether values are consistent with receiving a reward for targeted adverts or paying a fee for generic adverts. These ‘lower bound’ values, which are based on a combination of conservative approaches to valuation, are our preferred estimates on the basis that they are informed, conservative, and appropriately account for the fact that the CRR will not necessarily provide value to those that have already reviewed their settings.

As discussed in Appendix D, we find that our preferred estimates of the value of the CRR are not out of line with the values reported in the literature on privacy valuation.

### Aggregation

We estimate the total value of the CRR for the UK population of internet users based on our final estimates of the mean value (Table 12, approach (5))<sup>29</sup> and using the following data:

- ‘around 96% of UK internet users access at least one Google site each month. Facebook’s reach is around 87%.’ (CMA 2020a, p. 47).
- According to ONS estimates, there are 49,041k UK internet users aged 16+.<sup>30</sup>
- According to ONS estimates, the UK population aged 16–17 was 1,425,538 in 2019.<sup>31</sup>

Subtracting the total population aged 16–17 from the total number of UK internet users, as our survey covered the age group 18+, and multiplying by the above proportions of Google/Facebook users yields the following estimates:

- total number of UK Facebook users aged 18+: 41,425k;
- total number of UK Google users aged 18+: 45,711k.

Table 13 shows our estimates of the aggregate value of the CRR based on the above population totals of Facebook and Google users.

**Table 13:** Total value of the CRR to UK users of Facebook and Google (£/year)

Fee scenario		Reward scenario	
Uninformed choices	Informed choices	Uninformed choices	Informed choices
£527m [£434m–£620m]	£1,136m [£989m–£1,283m]	£1,761m [£1,441m–£2,082m]	£4,209m [£3,409m–£5,009m]

All estimates refer to a CRR providing a dichotomous choice between ‘Full data sharing’ (and receiving targeted adverts) and ‘No data sharing’ (and receiving ‘Generic’ adverts). All values are in £/year. 95% Confidence intervals, shown in parentheses, are based on Krinsky-Robb confidence intervals for the mean CRR valuation (see Table 12).

28 For these individuals, our argument that the CRR will expand the ‘choice set’ of Facebook/Google users does not hold, or it holds only to a lesser extent.

29 We do not use distinct CRR values for Facebook and Google users, as differences in valuations between these two groups are not statistically significant for any treatment (fee/reward; uninformed/informed).

30 ‘Used internet in the last 3 months’; reference period: January-March 2020; population: UK. <https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2020>

31 <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/adhocs/12663populationestimatesanddeathsbyingleyearofageforenglandandwalesandtheuk1961to2019>

Our recommended valuation of the CRR amounts to £1,136m per year. This estimate is based on the informed fee scenario. While uninformed choices provide a point of comparison for assessing the impact of information on the valuation of data privacy, the informed choices are those relevant for the valuation of the CRR, given that the complementary ‘Fairness by design’ intervention is aimed at ‘maximising users’ ability to make informed choices about the use of their personal data’ (CMA, 2020a). Our recommended valuation is a conservative estimate in several respects:

- While the ‘Fee’ scenario yields lower valuations compared to the ‘Reward’ scenario (an example of the WTA-WTP gap), the scenario described by the CMA (2020a), where platforms may ‘offer incentives to encourage consumers to agree to accept personalised advertising’ is effectively a ‘Reward’ scenario.
- We assume that the CRR valuation of those who have reviewed their data settings is zero. These users may, however, experience a welfare gain from a more transparent, more accessible, and wider choice between data sharing options following implementation of the CRR.
- We exclude users aged under 18 from our calculation of the total value of the CRR, effectively assigning a zero valuation for this user segment. Moreover, our segmentation analysis indicates that younger users place a higher value on the CRR than users aged 65+, suggesting that users’ valuation of the CRR may increase in the future.
- The survey sample was obtained from an online panel who, on account of them having signed up to reveal information about themselves for small amounts of money, are potentially more likely than the general population of users to be willing to share their data.

### Segmentation Analysis

For our segmentation analysis we used the same approach as for our main estimates. We ran separate econometric models for each segment and complement (to test for statistical differences) and used the parameter estimates to calculate segment-level mean CRR valuations as a weighted average of zero for those who had reviewed their data settings and the valuation of those who (a) had not reviewed their data settings<sup>32</sup> and (b) were not among the fastest 10% of participants and (c) were ‘Very sure’ or ‘Somewhat sure’ of their choices. For reasons given above under ‘Aggregation’, we focus on informed choices made in the ‘Fee’ scenario.

Table 14 provides a list of the segments covered in the analysis. Statistically significant differences in valuations across segments are shown in Table 15, higher (lower) values being highlighted in green (red). Each segment’s mean CRR valuation was compared against the valuation of the complement segment ‘Other’ (e.g., participants aged 18-29 vs participants aged 30 or above), and statistical significance of the differences (at the 5% level of significance) was assessed via t-tests.

32 Or, having reviewed the settings, found that they were not sure how to change them.

**Table 14:** User segments

Variable	Segments
Service use	{Facebook user; Google user}; Facebook and Instagram used {once a month or more often; never or less than once a month}
Highest frequency of use (across services)	{More than once a day; Once a day or less}
Aware of data collection by [Facebook/Google]	{Yes; No}
Targeted adverts on [Facebook/Google] are useful	{‘Very often’ or ‘Fairly often’; ‘Not very often’ or ‘Never’}
Level of comfort sharing personal data with [Facebook/Google]	{5 or lower; 6 or higher}
Ad blockers set up on any device	{Yes; No}
Awareness of Apple’s privacy changes in April 2021	{Yes; No}
Gender	{Female; Male}
Age	{18-29; 30-64; 65+}
Social grade	{A/B/C1; C2/D/E}; {A/B; C1/C2; D/E}
Education	{GCSE, GCE, or no formal qualification; Higher education below degree or A-Level; Undergraduate degree or higher}
Annual household income	{below £20,000; £20,000-£40,000; above £40,000}
Ethnicity	{White British and White Irish; Ethnic minorities (including White minorities) <sup>(1)</sup> }
Region	{London; South East; South West; East of England; East Midlands; West Midlands; Yorkshire and The Humber; North East; North West; Wales; Scotland; Northern Ireland}

(1) Including 37 who preferred not to answer

**Table 15:** Statistically significant differences in the valuation of the CRR across user segments

Segment	Mean value of CRR in informed fee scenario (£/user/month)	
<b>All participants</b>	<b>£1.09</b>	
Level of comfort sharing personal data: 5 or lower	£1.22	↑
Level of comfort sharing personal data: 6 or higher	£0.78	↓
Female	£1.23	↑
Male	£0.91	↓
Age 18-29	£1.17	↑
Age 30-64	£1.15	↑
Age 65+	£0.81	↓

Weighted average of the CRR valuation of those who had not reviewed their data settings and zero for those who had reviewed their settings (approach (5) in Table 12). Statistical significance was assessed via t-tests using the standard deviation of the simulated distribution of mean CRR values generated via Krinsky-Robb simulations as explained in Appendix C.

The CRR valuation of those 70% of participants whose level of comfort sharing data with Facebook/Google was 5 or lower (on a 1 to 10) scale was over 50% higher than for those who were more comfortable sharing data. Women had a higher valuation than men. Those aged 65 or above had a lower valuation than younger age groups.

The finding that those who are less comfortable sharing data have a considerably higher valuation of the CRR makes intuitive sense, and it is an indication that the survey has produced valid evidence about users' valuation of the CRR. This finding also helps to explain why some of the associations that one would expect to find in an explanatory model of differences in CRR valuations, such as higher valuations at higher incomes 'holding all else constant', do not translate into statistically significant segment differences. Those in higher income bands tended to have higher levels of comfort sharing data. This suggests opposite effects of these factors on the valuation of the CRR balancing out as an explanation for the absence of statistically significant differences in CRR valuations across income segments. The distributions of comfort levels were very similar for Facebook and Google users, as shown in Figure 17. This helps to explain why we have not found any statistically significant differences in valuations between Facebook and Google users.

Similarly, the proportion of those who had reviewed their data settings was substantially higher among those who had ad blockers. This explains why we do not find significant differences in valuations between those who had/did not have ad blockers, as higher valuations by the former are offset by a greater proportion of imputed zero CRR valuations by those who have ad blockers following our conservative 'lower bound' approach as set out above.

#### 5.4 Choices Between Data Sharing Alternatives

In addition to the binary contingent valuation questions examined above, the survey also included a final question offering different alternatives depending on the type of personal data that would be shared.

Table 16 and Table 17 show the proportions of choices for the different data sharing levels by payment vehicle (fee/reward) and service (Facebook/Google). Around 80% chose 'No data sharing' or 'Full data sharing', while the remaining 20% chose one of the intermediate options. This holds true across payment vehicles and services. The choices made in the final question largely conform to the overall pattern of choices in the binary contingent valuation questions:

- Overall, the proportion choosing 'Full data sharing' is over twice as high as the proportion choosing 'No data sharing' (55% vs 24%). Similarly, there were around two participants choosing 'Targeted' adverts for every participant who chose 'Generic' adverts in the first informed choice exercise involving payment of a fee/reward payment<sup>33</sup>.
- The proportions of Facebook and Google users choosing each data sharing option were very similar, as were the choices made by these two groups of participants in the binary contingent valuation questions.

Looking at intermediate data sharing options, the choice share of '1<sup>st</sup> party data only' was somewhat higher than those of the remaining two options which had approximately equal choice shares.

33 The relative proportion of those choosing 'Targeted'/'Full data sharing' to those choosing 'Generic'/'No data sharing' was higher (lower) in the final questions than in the binary contingent valuation question in the fee (reward) treatment (see Table 16 vs Table 10).

**Table 16:** Choice of data sharing level by payment vehicle

	All (Base: 4,014)	Fee (Base: 2,000)	Reward (Base: 2,014)
No data sharing	24%	11%	38%
1 <sup>st</sup> party data only	9%	9%	9%
1 <sup>st</sup> party data and 3 <sup>rd</sup> party online tracking	6%	6%	6%
1 <sup>st</sup> party data and 3 <sup>rd</sup> party customer lists	6%	5%	7%
Full data sharing	55%	70%	40%
<b>Total</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>

Note: Percentages may not add up to 100% due to rounding

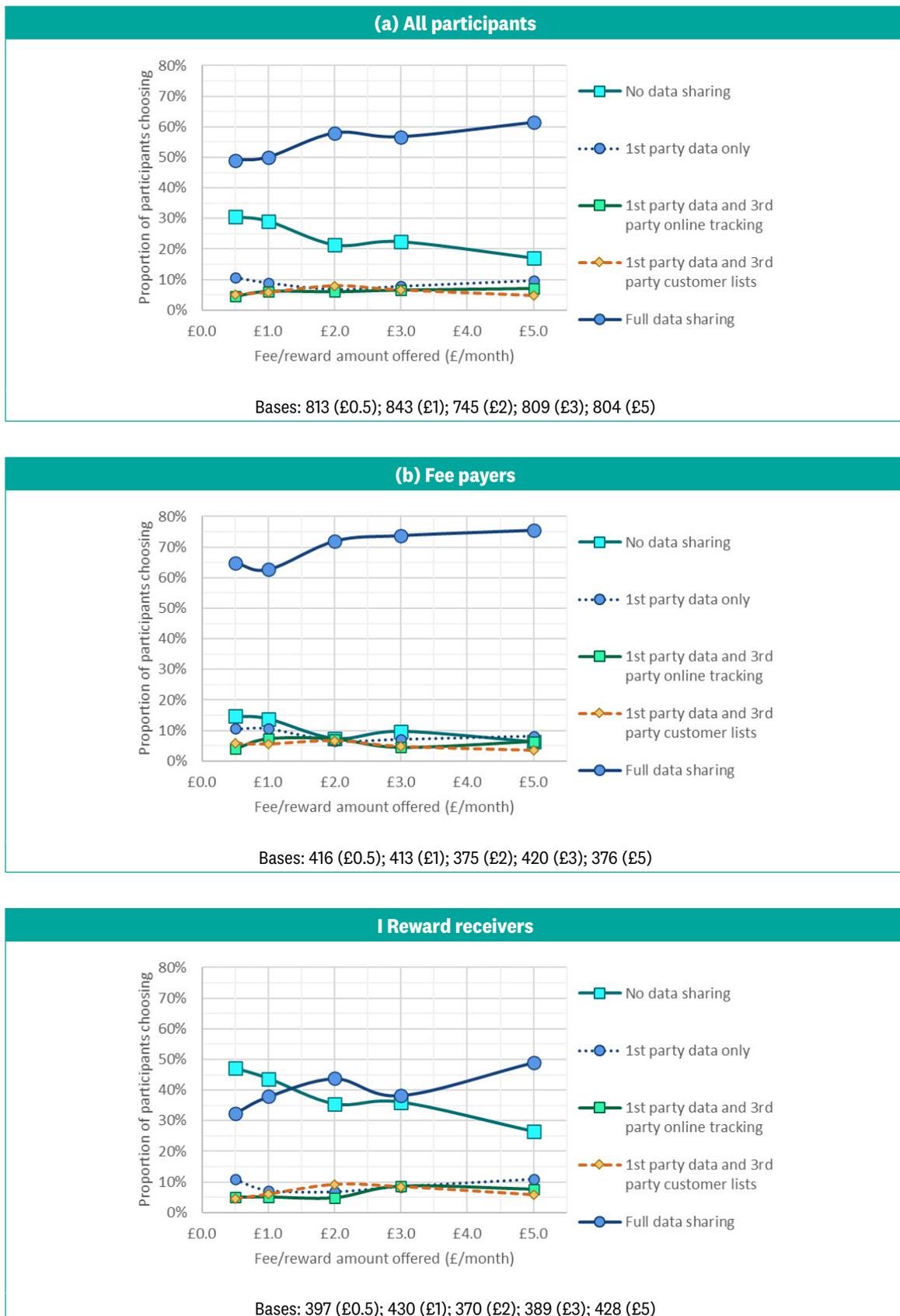
**Table 17:** Choice of data sharing level by service

	All (Base: 4,014)	Facebook (Base: 2,005)	Google (Base: 2,009)
No data sharing	24%	23%	25%
1 <sup>st</sup> party data only	9%	10%	8%
1 <sup>st</sup> party data and 3 <sup>rd</sup> party online tracking	6%	7%	6%
1 <sup>st</sup> party data and 3 <sup>rd</sup> party customer lists	6%	6%	6%
Full data sharing	55%	55%	55%
<b>Total</b>	<b>100%</b>	<b>100%</b>	<b>100%</b>

Note: Percentages may not add up to 100% due to rounding

Figure 30 shows how the proportion of choices of the different data sharing levels varied by the amount of fee/reward.

**Figure 30:** Choice of data sharing alternatives



The charts show that participants' choices were sensitive to cost, especially for the 'Not data sharing' and 'Full data sharing' options. As expected, the proportions choosing 'No data sharing' tended to decrease with the base reward/fee amounts, while the proportions choosing 'Full data sharing' increased with the base reward/fee amounts.

Overall, these results do not provide strong evidence to suggest that participants have substantial differences in preferences with respect to the different types of data that are potentially shared for the purposes of targeted advertising. This finding is consistent with the findings in Section 4.5, which showed that the level of comfort was broadly similar between third party online tracking and third party customer lists.

## 6. Conclusions

The principal objective of the present study was to obtain user valuations of the CMA's proposed CRR. The estimated valuations vary from £0.50 per user per month to £4.03 per user per month depending on whether the valuation is based on an informed or an uninformed question and whether values are consistent with receiving a reward for targeted adverts or paying a fee for generic adverts.

Combining the most relevant scenario for valuation of the CRR (i.e., informed choices) and the most conservative approach to valuation (i.e., the 'fee treatment') yields an estimate of £1.09 per user per month.

The latter value implies a total valuation of the CRR by UK Facebook and Google users of £1,136m per year.

This is a conservative estimate for the following reasons:

- It is based on the lower 'Fee' scenario whereas the scenario envisaged by the CMA (2020a), where platforms may 'offer incentives to encourage consumers to agree to accept personalised advertising', is effectively a 'Reward' scenario.
- It is assumed that the CRR valuation of those who have reviewed their data settings is zero. These users may, however, experience a welfare gain from a more transparent and accessible choice between data sharing options following implementation of the CRR.
- We exclude users aged under 18 from our calculation of the total value of the CRR, effectively assigning a zero valuation for this user segment. Moreover, our segmentation analysis indicates that younger users place a higher value on the CRR than users aged 65+, suggesting that users' valuation of the CRR may increase in the future.
- The survey sample was obtained from an online panel who, on account of them having signed up to reveal information about themselves for small amounts of money, are potentially more likely than the general population of users to be willing to share their data.

There are several indications supporting the validity and reliability of the valuations:

- An extensive development and testing programme found that the core choices were considered plausible by participants, and that participants had no problems answering these questions appropriately.
- Values were obtained from a large national sample of more than 4,000 Google and Facebook users.
- Valuations vary in line with expectations:
  - Informed' valuations were higher than 'Uninformed' valuations.
  - Valuations were higher when associated with receiving a reward for targeted adverts rather than paying a fee for generic adverts.
  - Those who were not comfortable sharing their personal data expressed a higher valuation.
- Values are consistent with other estimates reported in the literature on privacy valuation. A comparison of values in the literature across the dimensions of privacy that are relevant to

our study and comparable to the units of valuation of our study, show that our recommended range for user valuations of the CRR of £0.50-£4.03 per user per month lies within the range of values of online privacy reported in the literature, i.e., £0.41 to £6.40 per user per month.

There are hence good grounds for considering the valuations of the CRR to be robust and reliable for users of Facebook and Google in the UK.

## References

- Acquisti, Alessandro (2004) "Privacy in Electronic Commerce and the Economics of Immediate Gratification." In EC '04: Proceedings of the 5<sup>th</sup> ACM Conference on Electronic Commerce, 21–29. New York: Association for Computing Machinery.
- Acquisti, Alessandro and Grossklags, Jens (2005) "Privacy and Rationality in Individual Decision Making." *IEEE Security and Privacy*, 3(1), 26–33.
- Acquisti, Alessandro, and Hal R. Varian (2005) "Conditioning Prices on Purchase History." *Marketing Science* 24 (3): 367–81.
- Acquisti, Alessandro, and Ralph Gross (2006) "Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook". *Privacy Enhancing Techniques* 4258: 36–58.
- Acquisti, Alessandro, Leslie K. John, and George Loewenstein (2013) "What Is Privacy Worth?" *Journal of Legal Studies* 42 (2): 249–74.
- Acquisti, Alessandro, Laura Brandimarte, and George Loewenstein (2015) "Privacy and Human Behaviour in the Age of Information." *Science* 347 (6221): 509–14.
- Acquisti, Alessandro, Curtis Taylor and Liad Wagman (2016) "The Economics of Privacy." *Journal of Economic Literature* 54(2): 442–492.
- Acquisti, Alessandro et. Al (2017) "Nudges for Privacy and Security: Understanding and Assisting Users' Choices Online". *ACM Computing Surveys* 50(3) Article 44.
- Adjerid, Idris, Alessandro Acquisti, Laura Brandimarte, and George Loewenstein (2013) "Sleights of Privacy: Framing, Disclosures, and the Limits of Transparency." In *Proceedings of the Ninth Symposium on Usable Privacy and Security*, 1–11. New York: Association for Computing Machinery.
- Alberini, Anna, Barbara Kanninen, and Richard T. Carson (1997) "Modeling Response Incentive Effects in Dichotomous Choice Contingent Valuation Data". *Land Economics* 73(3): 309–324.
- Berendt, Bettina, Oliver Gunther, and Sarah Spiekermann (2005) "Privacy in e-Commerce: Stated Preferences vs. Actual Behaviour." *Communications of the ACM* 48(4): 101–06.
- Beresford, Alastair R. & Kübler, Dorothea & Preibusch, Sören (2012) "Unwillingness to pay for privacy: A field experiment." *Economics Letters*, Elsevier, 117(1): 25–27.
- Bergemann, Dirk, and Alessandro Bonatti (2015) "Selling Cookies." *American Economic Journal: Microeconomics* 7(3): 259–94.
- Berthold, Stefan, and Rainer Böhme (2010) "Valuating Privacy with Option Pricing Theory." In *Economics of Information Security and Privacy*, edited by Tyler Moore, David Pym, and Christos Ioannidis, 187–209. New York: Springer.

- Board, Simon, and Jay Lu (2015) "Competitive Information Disclosure in Search Markets." Unpublished.
- Brandimarte, Laura, Alessandro Acquisti, and George Loewenstein (2013) "Misplaced Confidences: Privacy and the Control Paradox." *Social Psychological and Personality Science* 4 (3): 340–47.
- Calzolari, Giacomo, and Alessandro Pavan (2006) "On the Optimality of Privacy in Sequential Contracting." *Journal of Economic Theory* 130 (1): 168–204.
- Campbell, James, Avi Goldfarb, and Catherine Tucker (2015) "Privacy Regulation and Market Structure." *Journal of Economics and Management Strategy* 24 (1): 47–73.
- Carson, R. and T. Groves (2007) Incentive and informational properties of preference questions. *Environmental and Resource Economics*, 37: 181-210.
- Chellapa, Ramnath K., and Raymond G. Sin (2005) "Personalization versus Privacy: An Empirical Examination of the Online Consumers' Dilemma". *Information Technology and Management* 6:181–202.
- CMA (2020a) Online platforms and digital advertising: Market study final report.
- CMA (2020b) Online platforms and digital advertising: Market study final report : Appendix L: Summary of research on consumers' attitudes and behaviour.
- Conitzer, Vincent, Curtis R. Taylor, and Liad Wagman (2012) "Hide and Seek: Costly Consumer Privacy in a Market with Repeat Purchases." *Marketing Science* 31 (2): 277–92.
- Cvrcek, Dan, Marek Kumpost, Vashek Matyas, and George Danezis (2006) "A Study on the Value of Location Privacy" Pp. 109–18 in *Proceedings of the 5<sup>th</sup> ACM Workshop on Privacy in the Electronic Society (WPES '06)*, edited by Ari Juels and Marianne Winslett. New York: Association for Computing Machinery.
- Day, Brett (2007) "Distribution-free estimation with interval-censored contingent valuation data: troubles with Turnbull?". *Environmental Resource Economics* 37: 777–95.
- De Corniere, Alexandre (2016) "Search Advertising." *American Economic Journal: Microeconomics* 8(3):156-88.
- Egelman, S., Porter Felt, A., and Wagner D. (2013) 'Choice architecture and smartphone privacy: There's a price for that.' In *The Economics of the Information Society*. Springer, 211-236.
- Friedman, Eric J., and Paul Resnick (2001) "The Social Cost of Cheap Pseudonyms." *Journal of Economics and Management Strategy* 10 (2): 173–99.
- Gentleman R., and Alain Vandal (2021). *Icens: NPMLE for Censored and Truncated Data*. R package version 1.64.0.
- Goldstein, D.G., Johnson, E.J., Herrman, A., and Heitmann, M. (2008) 'Nudge your customers towards better choices.' *Harvard Business Review*. 86, 12, 99-105.
- Grossklags, J., and Acquisti, A. (2007) 'When 25 cents is too much: An experiment on willingness-to-sell and willingness-to-protect personal information. In *Proceedings of the Workshop on the Economics of Information Security (WEIS'07)*, 1-22.

- Haab, Tim and McConnell, Kenneth E. (2002) *Valuing Environmental and Natural Resources*, Edward Elgar Publishing.
- Hann, Il-Horn, Kai-Lung Hui, Sang-Yong Tom Lee, and Ivan P. L. Png. (2007) “Overcoming Information Privacy Concerns: An Information Processing Theory Approach”. *Journal of Management Information Systems* 24:13–42.
- Hagiu, Andrei, and Bruno Jullien (2011) “Why Do Intermediaries Divert Search?” *RAND Journal of Economics* 42 (2): 337–62.
- Hirshleifer, Jack (1980) “Privacy: Its Origin, Function, and Future”. *Journal of Legal Studies* 9:649–66.
- Hoofnagle, Chris Jay, and Jennifer M. Urban (2014) “Alan Westin’s Privacy Homo Economicus.” *Wake Forest Law Review* 49 (261).
- Huberman, Bernardo A., Eytan Adar, and Leslie Fine (2005) “Valuating Privacy”. *IEEE Security and Privacy* 3:22–25.
- Hui, Kai-Lung, H-H. Teo, and Sang-Yong Lee (2007) “The Value of Privacy Assurance: An Exploratory Field Experiment”. *MIS Quarterly* 31:19–33.
- Jentzsch, Nicola, Soren Preibusch, and Andreas Harasser (2012) “Study on Monetising Privacy: An Economic Model for Pricing Personal Information.” Berlin: European Network and Information Security Agency.
- John, Leslie K., Alessandro Acquisti, and George Loewenstein (2011) “Strangers on a Plane: ContextDependent Willingness to Divulge Sensitive Information.” *Journal of Consumer Research* 37 (5): 858–73.
- Johnson, E.J. (2012) ‘Beyond Nudges: Tools of a Choice Architecture.’ *Marketing Letters*, Vol 23, 487–504.
- Kelley, P.G., Cesca, L., Bresee, J., and Cranor, L.F. (2010). ‘Standardizing privacy notices: An online study of the nutrition label approach’, In *Proceedings of the Conference on Human Factors in Computing Systems (CHI’10)*, ACM,1573-1582.
- Lai, Y-L., and Hui, K-L. (2004) ‘The Opt-in and Opt-Out debate: is it really necessary?’ Available at: [http://isr.uci.edu/pep06/papers/PEP06\\_LaiHui.pdf](http://isr.uci.edu/pep06/papers/PEP06_LaiHui.pdf).
- Laudon, Kenneth C. (1997) “Extensions to the Theory of Markets and Privacy: Mechanics of Pricing Information.” New York University Stern School of Business Working Paper IS-97-4.
- Leon, P.G., Ur, B., Balebako, R., Cranor, L.F., Shay, R., Wang, Y. (2011), ‘Why Johnny Can’t Opt-Out: A Usability Evaluation of Tools to Limit Online Behavioural Advertising.’ In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 589-598.
- Mathur, A., Acar, G., Friedman, M.J., Lucherini, E., Mayer, J., Chetty, M., and Narayanan, A. (2019) ‘Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites.’ In *Proceedings of the ACM on Human-Computer Interaction Volume 3, Issue CSCW*, Article 81.
- McDonald, Aleecia, and Lorrie Faith Cranor (2010) “Beliefs and Behaviours: Internet Users’ Understanding of Behavioural Advertising.” Paper presented at the 2010 TPRC—43<sup>rd</sup> Research Conference on Communication, Information and Internet Policy, Arlington, VA, October 2 2010.

- Nakatani, Tomoaki, Hideo Aizaki, and Kazuo Sato (2020). Dcchoice: An R Package for Analyzing Dichotomous Choice Contingent Valuation Data. R package version 0.0.17.
- Norberg, Patricia A., Daniel R. Horne, and David A. Horne (2007) “The Privacy Paradox: Personal Information Disclosure Intentions versus Behaviors.” *Journal of Consumer Affairs* 41 (1): 100–126.
- Norwegian Consumer Council (2018). ‘Deceived by Design.
- Norwegian Consumer Council (2018). Every Step You Take.
- OECD (2018) Maintaining competitive conditions in the era of digitalisation: Report to G-20 Finance Ministers and Central Bank Governors, July 2018 (accessed at <https://www.oecd.org/g20/Maintaining-competitive-conditions-in-era-of-digitalisation-OECD.pdf>).
- OECD (2019) Regulatory effectiveness in the era of digitalisation (accessed at <https://www.oecd.org/gov/regulatory-policy/Regulatory-effectiveness-in-the-era-of-digitalisation.pdf>).
- Olejnik, Lukasz, Tran Minh-Dung, and Claude Castelluccia (2014) “Selling Off Privacy at Auction.” Paper presented at ISOC Network and Distributed System Security Symposium, San Diego, CA, February 23 2014.
- ONS (2020) Internet use in the UK; annual estimates by age, sex, disability and geographical location. (accessed at: <https://www.ons.gov.uk/businessindustryandtrade/itandinternetindustry/bulletins/internetusers/2020>).
- ONS (2020) Population estimates and deaths by single year of age for England and Wales and the UK, 1961 to 2019’. (accessed at: <https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/adhocvs/12663/populationestimatesanddeathsbysingleyearofageforenglandandwalesandtheuk1961to2019>).
- Png, Ivan P. L. (2007) “On the Value of Privacy from Telemarketing: Evidence from the “Do Not Call” Registry. Working paper. National University of Singapore, School of Business.
- Posner, Richard A. (1978) “The Right of Privacy” *Georgia Law Review* 12:393–422.
- Posner, Richard A. (1981) “The Economics of Privacy” *American Economic Review* 71:405–9.
- Prince, Jeffrey and Wallsten, S. (2020) “How Much is Privacy Worth Around the World and Across Platforms?” TPRC48: The 48<sup>th</sup> Research Conference on Communication, Information and Internet Policy.
- Quinn, K. (2016) ‘Why We Share: A Uses and Gratifications Approach to Privacy Regulation in Social Media Use.’ *Journal of Broadcasting & Electronic Media* 60(1), 61–86.
- R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Rainie, Lee, Sara Kiesler, Ruogu Kang, and Mary Madden (2013) “Anonymity, Privacy, and Security Online.” Washington, DC: Pew Research Centre.
- Ready, Richard C., Patricia A. Champ, and Jennifer L. Lawton (2010) “Using Respondent Uncertainty to Mitigate Hypothetical Bias in a Stated Choice Experiment”, *Land Economics* 86 (2), 363–81.

- Rose, Ellen A. (2005) "Data Users versus Data Subjects: Are Consumers Willing to Pay for Property Rights to Personal Information?" Proceedings of the 38<sup>th</sup> Annual Hawaii International Conference on System Sciences. Washington, D.C.: IEEE Computer Society Press. Doi:10.1109/HICSS.2005.184.
- Saetra, H. (2020) "Privacy as an aggregate public good" *Technology in Society*, Volume 63, 101422.
- Savage, Scott and Waldman, Donald M., (2013) "The Value of Online Privacy" Available at SSRN:<https://ssrn.com/abstract=2341311> or <http://dx.doi.org/10.2139/ssrn.2341311>.
- Shklovski, I., Mainwaring, S.D., Skúladóttir, H.H., and Borgthorsson, H. (2014). 'Leakiness and Creepiness in App Space: Perceptions of Privacy and Mobile App Use.' In Proceedings of the Conference on Human factors in Computing Systems, (CHI'14): <https://doi.org/10.1145/2556288.2557421>.
- Spiekermann, Sarah, Jen Grossklags, and Bettina Berendt (2001) "E-Privacy in Second Generation E-Commerce: Privacy Preferences versus Actual Behavior". Pp. 38–47 in Proceedings of the Third ACM Conference on Electronic Commerce, edited by Michael P. Wellman and Yoav Shoham. New York: Association for Computing Machinery.
- Statista (2021) <https://www.statista.com/statistics/1228227/online-privacy-concerns-uk/>.
- Stigler, George J. (1980) "An Introduction to Privacy in Economics and Politics". *Journal of Legal Studies* 9:623–44.
- Swire, Peter P., and Robert E. Litan. (1998) "None of Your Business: World Data Flows, Electronic Commerce, and the European Privacy Directive". Washington, DC: Brookings Institution Press.
- Taylor, Curtis R. (2004) "Consumer Privacy and the Market for Customer Information." *RAND Journal of Economics* 35 (4): 631–50.
- Tedeschi, Bob. (2002) Everybody Talks about Online Privacy, but Few Do Anything about It. *New York Times*, June 3 2002.
- Train, Kenneth (2003) *Discrete Choice Methods with Simulation*. Cambridge University Press.
- Tsai, Janice Y., S. Egelman, L. Cranor, and Alessandro Acquisti. (2011) "The Effect of Online Privacy Information on Purchasing Behavior: An Experimental Study". *Information Systems Research* 22:254–68.
- Turnbull, Bruce W. (1976) "The Empirical Distribution Function with Arbitrarily Grouped, Censored and Truncated Data". *Journal of the Royal Statistical Society. Series B (Methodological)* 38 (3): 290–5.
- Turow, Joseph, Jennifer King, Chris Jay Hoofnagle, Amy Bleakley, and Michael Hennessy. (2009) "Americans Reject Tailored Advertising and Three Activities That Enable It." [http://repository.upenn.edu/cgi/viewcontent.cgi?article=1138&context=asc\\_papers](http://repository.upenn.edu/cgi/viewcontent.cgi?article=1138&context=asc_papers).
- Varian, Hal R. (1997) "Economics Aspects of Personal Privacy." In *Privacy and Self-Regulation in the Information Age*. Washington, DC: US Department of Commerce, National Telecommunications and Information Administration.
- Varian, Hal R., Fredrik Wallenberg, and Glenn Woroch. (2005) "The Demographics of the Do-Not-Call List." *IEEE Security and Privacy* 3 (1): 34–39.

Villas-Boas, J. Miguel. (2004) "Price Cycles in Markets with Customer Recognition." *RAND Journal of Economics* 35(3): 486–501.

Wathieu, Luc, and Allan Friedman. (2007) "An Empirical Approach to Understanding Privacy Valuation". Working Paper No. 07-075. Harvard Business School, Cambridge, Mass.

Winegar, A.G. & Sunstein, C.R., (2019). 'How much is data privacy worth? A preliminary investigation.' *Journal of Consumer Policy* 42: 425–440.

Zafeiropoulou, A.M., Millard, D.E., Webber, C., and O'Hara, K. (2013) 'Unpicking the privacy paradox: can structuration theory help to explain location-based privacy decisions? *ACM Web Science*, Paris, France. <http://dx.doi.org/10.1145/2464464.2464503>

# Appendix A – Main Survey Questionnaire



3460  
Which? Quant Questionnaire V19

SYSTEM INFORMATION :

Date:

Time interview started:

## Which? Personalised Advertising Research

Thank you very much for agreeing to complete this on-line survey which is being conducted by Accent.

We are interested in understanding attitudes to advertising that you might see when using search engines and social media and the questionnaire should take about 15 minutes.

The research is being done on behalf of Which? So they can better understand the advertising that people see online and how they feel about it. This work could potentially be used to advise the Government on legislation that they might bring in around online advertising.

Any answer you give will be treated in confidence in accordance with the Code of Conduct of the Market Research Society. If you would like to confirm Accent's credentials type Accent in the search box at: <https://www.mrs.org.uk/researchbuyersguide>.

**IF MOBILE DEVICE SHOW:** This survey is best undertaken on a tablet or a PC. If you do use a smartphone you can switch between desktop mode and mobile mode at any time by clicking the button at the bottom of the screen.

We will just ask you a couple of questions to check that you are eligible to take part in this research. For convenience you can stop and return to complete the questionnaire as many times as you wish, although once submitted you will not be able to enter again. The date for completion is 25 May 2021

### Scoping questions

Q1. Any data collected over the course of this interview that could be used to identify you, such as your name, address, or other contact details, will be held securely and will not be shared with any third party unless you give permission (or unless we are legally required to do so). Our privacy statement is available at <https://www.accent-mr.com/privacy-policy/>.

Do you agree to proceeding with the survey on this basis?

Yes

No THANK AND CLOSE

---

Q2. What gender do you identify as?

Female

Male

Other

Prefer not to say

DP: CHECK QUOTA

---

Q3. How old are you?

18-29

30-44

- 35-44
- 45-64
- 65 or over
- Prefer not to say

DP: CHECK QUOTA

---

Q4. Please select from the list below the region you're currently living in.

- North East
- North West
- Yorkshire and The Humber
- East Midlands
- West Midlands
- East of England
- London
- South East
- South West
- Wales
- Scotland
- Northern Ireland
- Prefer not to say **THANK AND CLOSE**

DP: CHECK QUOTA

---

Q5. **APPROX. SEG:** How would you describe the occupation of the chief income earner in your household?

- Senior managerial or professional
- Intermediate managerial, administrative or professional
- Supervisor; clerical; junior managerial, administrative or professional
- Manual worker (with industry qualifications)
- Manual worker (with no qualifications)
- Unemployed due to ill health
- Unemployed or not working for another reason
- Retired
- Student
- Prefer not to say

Q6. **IF Q5=8 (RETIRED), ASK, ELSE SKIP:** Does the chief income earner have a state pension, a private pension or both?

- State only
- Private only
- Both

IFQ6=1 (STATE ONLY), GO TO Q8

---

Q7. **IF Q6= PRIVATE OR BOTH, ASK, ELSE SKIP:** How would you describe the chief income earner's occupation before retirement?

- Senior managerial or professional
- Intermediate managerial, administrative or professional
- Supervisor; clerical; junior managerial, administrative or professional
- Manual worker (with industry qualifications)
- Manual worker (with no qualifications)
- None of these

Q7b. **SEG: CODE AS FOLLOWS:**

---

IF Q5= 1 or 2; SEG = AB  
 IF Q5= 3; SEG = C1  
 IF Q5= 4; SEG = C2  
 IF Q5= 5; SEG = DE  
 IF Q5= 6; SEG = DE  
 IF Q5= 7; SEG = DE  
 IF Q5= 9; SEG = C1

IF Q5= 8 and 0 = State only; SEG = DE

IF Q5= 8 and 0= Private only OR Both and 0= 1; SEG = AB  
 IF Q5= 8 and 0= Private only OR Both and 0= 2; SEG = AB  
 IF Q5= 8 and 0= Private only OR Both and 0= 3; SEG = C1  
 IF Q5= 8 and 0= Private only OR Both and 0= 4; SEG = C2  
 IF Q5= 8 and 0= Private only OR Both and 0= 5; SEG = DE  
 IF Q5= 8 and 0= Private only OR Both and 0= 6; SEG = DE

IF Q5= 10; SEG = Not stated

DP: CHECK QUOTAS

Q8. In the past three months, how often have you used any of the following Google or Facebook services on any device:

RANDOMISE ORDER

	More than once a day	Once a day	At least once a week	At least once a month	Less than once a month	Never use this service
Facebook (the main app rather than messenger)						
Instagram						
Google Search – a search engine that allows you to search the web (similar to Bing and Yahoo)						
Google Chrome – a web browser (similar to Firefox, Internet Explorer, Safari)						
Google Maps						
YouTube						

SCREEN OUT IF “Never used this service” FOR ALL SERVICES or if just Instagram is selected  
 Please also screen out if less than once a month for all services or if just Instagram is selected

Random allocation between Facebook and Google unless only use one. Representative sample by Census going into Q8.

Please exclude Instagram from the Facebook grouping. i.e. if they just use Instagram and nothing else they should close

## Main questionnaire

Thank you, I can confirm you are in scope for the survey. The questionnaire will take no more than 15 minutes.

For convenience you can stop and return to complete the questionnaire as many times as you wish, although once submitted you will not be able to enter again.

Throughout the survey there are a number of information buttons ⓘ that you can click on for further information.

Q8b Which of the following devices do you use?

- Smartphone
- Tablet
- Laptop
- Desk top computer

Q8c **FOR EACH DEVICE SELECTED AT Q8b:** Which of the following operating systems do you use? If you have more than one of each device please think about the device you use **most frequently** for **personal** use

- Apple (e.g. iOS, macOS)
- Android (e.g. Samsung....)
- Windows
- Other
- Don't know

### Targeted advertising awareness and preferences

Q9. Are you aware that **[Facebook/Google]** uses data that it collects about you to target adverts that it thinks are relevant to you? (This means you might see different adverts to another person).

- Yes
- No

Q10. How often do you find advertisements on **[Facebook/Google]** that are targeted (i.e. intended to be relevant) useful?

- Very often
- Fairly often
- Not very often
- Never
- Don't know

Q11. In general, how comfortable are you sharing your personal data with **[Facebook/Google]** for the purpose of targeted advertising? Please mark your answer on a scale of 1 to 10 where 1 is 'Not comfortable at all' and 10 is 'Completely comfortable'

- 1 Not comfortable at all
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10 Completely comfortable

Q12. Soon **[Facebook/Google]** may have to ask you if you want them to use your personal data to receive targeted adverts. If you didn't see targeted adverts, you would still see adverts, they would just be generic i.e. not personalised to you using your personal data.

Suppose you were given the following choice, which option would you prefer, A or B?

[RANDOM VARIABLE: OPTIONORDER = 1 OR 2]

IF OPTIONORDER=1, SHOW OPTION A="See targeted adverts"; OPTION B="See generic adverts"  
 IF OPTIONORDER=2, SHOW OPTION B="See targeted adverts"; OPTION A="See generic adverts"

<b>Option A</b> See targeted adverts	<b>Option B</b> See generic adverts
<input type="radio"/>	<input type="radio"/>

**Uninformed Choice**

BASECOST={£0.50, £1.00, £2.00, £3.00, £5.00} (Select one amount at random)  
 REWARDFEE= REWARD OR FEE AT RANDOM  
 IF REWARDFEE=FEE SKIP TO Q17

Q14a [REWARDFEE=REWARD ONLY] [Facebook/Google] may offer you an incentive to choose to receive targeted adverts. Which would you prefer, assuming the amount was of equal value?

A payment through your PayPal account  
 Amazon vouchers/gift card  
 Either  
 [CREATE REWARDTYPE=ANSWER]

Q14 Suppose you were given the following choice. Which option would you choose, A or B?

IF OPTIONORDER=1, SHOW OPTION A="See targeted adverts"; OPTION B="See generic adverts"  
 IF OPTIONORDER=2, SHOW OPTION B="See targeted adverts"; OPTION A="See generic adverts"

<b>Option A</b> See targeted adverts You will receive rewards worth BASECOST each month from [Facebook/Google] These rewards will be paid at the start of each month If you change your mind you will continue to receive targeted adverts until the end of the month but can then opt out from targeted adverts	<b>Option B</b> See generic adverts No reward will be paid to you
<input type="radio"/>	<input type="radio"/>

Q16a How sure are you that this is the choice you would make in practice if this choice was really offered to you?

Very sure  
 Somewhat sure  
 Not very sure  
 Not at all sure

CREATE BASECOST2  
 =2\*BASECOST IF Q14="Generic"  
 =0.5\*BASECOST IF Q14="Targeted"

Q16 Now suppose you were given the following choice. Imagine this is a different scenario and not linked to the choice you have just been shown. Which option would you choose, A or B?

[SHOW SAME CHOICE AS AT Q14 BUT WITH BASECOST REPLACED WITH BASECOST2]

SKIP TO 'How your data is collected'

Q17. [REWARDFEE=FEE ONLY] [Facebook/Google] may require you to pay a monthly fee to continue to use its service with generic adverts being shown rather than targeted adverts.

Suppose you were given the following choice. Which option would you choose, A or B?

IF OPTIONORDER=1, SHOW OPTION A="See targeted adverts"; OPTION B="See generic adverts"  
 IF OPTIONORDER=2, SHOW OPTION B="See targeted adverts"; OPTION A="See generic adverts"

Option A See targeted adverts	Option B See generic adverts
There would be no fee payable	You will pay <b>BASECOST</b> per month to continue using [Facebook/Google] without targeted adverts. This fee must be paid at the start of each month. If you change your mind you will continue to receive generic adverts until the end of the month but can then Opt-in to targeted adverts.

Q18. Why did you choose this option?

Q19a How sure are you that this is the choice you would make in practice if this choice was really offered to you?

- Very sure
- Somewhat sure
- Not very sure
- Not at all sure

DP: CREATE BASECOST2  
 =2\*BASECOST IF Q17="Generic"  
 =0.5\*BASECOST IF Q17="Targeted"

Q19. Now suppose you were given the following choice. Which option would you choose, A or B?

[SHOW SAME CHOICE AS AT Q17 BUT WITH BASECOST REPLACED WITH BASECOST2]

## How your data is collected

### Intro for Facebook

This next part is about your current settings and the personal information that Facebook currently has access to. This isn't what you share with family and friends or general public. This is related to how Facebook uses your data to make the ads that you see more relevant.

Facebook needs to collect information about you in order to decide which adverts will be most relevant to you.

**Intro for Google**

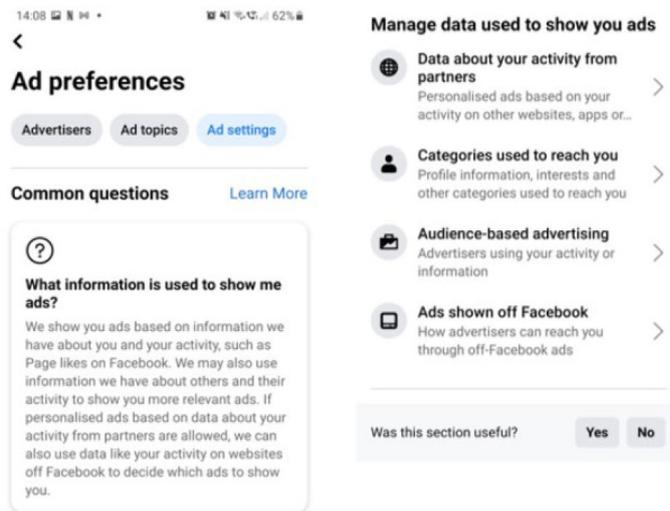
This next part is about your current settings and the personal information that Google currently has access to. This is related to how Google uses your data to make the ads that you see more relevant.

Google needs to collect information about you in order to decide which adverts will be most relevant to you.

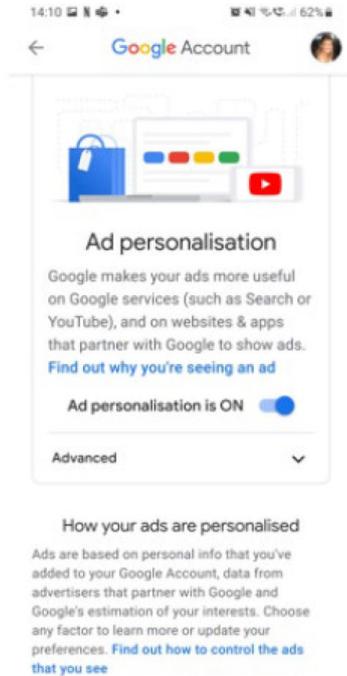
---

Q20. Were you aware that you could change your settings relating to your data and how [Facebook/Google] can use that information to inform the adverts that you see when you are using [Facebook/Google]?

FACEBOOK IMAGE:



GOOGLE IMAGE



- Yes, and I have changed my settings
- Yes, I have reviewed my settings but chose not to change them
- Yes, but I have never gone in and looked at them
- No, I wasn't aware
- Don't know

---

Q21. IF Q20="Yes, and I have changed my settings"] Why did you change your settings?

- I like to be able to choose how my data is used
- To see fewer adverts
- To see fewer targeted adverts
- I don't want to share my data
- Other (please write in)

---

Q22. IF Q20="Yes, and I have changed my settings"] What setting(s) did you change?

---

Q23. IF Q20="Yes, I have reviewed my settings but chose not to change them"] Why did you choose not to change your settings?

- Happy with them as they are
- I wasn't sure how to
- Other (please write in)

---

I'm now going to tell you about three of the ways [Facebook/Google] collects information and ask how you feel about each of them.

Slide 1 Facebook	Slide 1 Google
<p><b>First-party data collection</b></p> <p>One way Facebook collects information about you is by observing what you do on Facebook and Instagram (which Facebook also owns) and from your profile information. If you have accounts with both, information about what you do on Instagram will influence what you see on Facebook. Facebook then uses this data to try to make the adverts you see at Facebook more relevant.</p> <p>It collects information on things like:</p> <ul style="list-style-type: none"> <li>• what posts and pages you “like”</li> <li>• information you have added to your Facebook and Instagram profile</li> <li>• places that you “check into”</li> <li>• ads and other content that you watch or engage with (e.g. click on, videos that you watch).</li> </ul> <p>By collecting this information Facebook knows things like your age, gender, location and what your interests are. This is then used to target you with adverts that Facebook thinks you will be interested in.</p>	<p><b>First-party data collection</b></p> <p>Google collects information about you in these two ways.</p> <p>1. Through Google services (e.g. Google Maps, Google Chrome, YouTube, Google Search etc.)</p> <ul style="list-style-type: none"> <li>• information you provide when creating your Google account (e.g. name, contact details, gender, payment information).</li> <li>• information about apps, browsers and devices that you use to access or interact with Google services (e.g. the name of the browser you use, your operating system, the IP address of your computer etc.).</li> <li>• information about your online activity when you use Google services (e.g. the contents of your shopping basket, advertisements you see, pages you visit, your YouTube watch history etc.).</li> <li>• information about your location.</li> </ul> <p>2. Through Android mobile devices that use Google’s own operating system. Google collects information about you (such as your location data, your unique mobile device identifier (known as IMEI) etc.).</p> <p>By collecting this information Google knows things like your age, gender, location and what your interests are. This is then used to target you with adverts that Google thinks you will be interested in.</p>

**Q24. First party data collection**

How comfortable are you for **[Facebook/Google]** to collect first party data in order to inform targeted adverts? Please mark your answer on a scale of 1 to 10 where 1 is ‘Not comfortable at all’ and 10 is ‘Completely comfortable’

- 1 Not comfortable at all
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10 Completely comfortable

Slide 2 Facebook	Slide 2 Google
<p><b>Third-party: online tracking</b></p> <p>Facebook also collects data about you from other organisations i.e. other businesses or charities that have a relationship with Facebook. Facebook uses this data to try to make the adverts you see on Facebook more relevant.</p> <p>One of the ways they collect information from other organisations is to track your activity when you’re on websites and apps other than Facebook. They do this using tools like cookies, or if you log-in using your Facebook log-in.</p> <p>Information is sent back to Facebook about web pages you’ve looked at, whether you’ve purchased something or put something in your online basket.</p>	<p><b>Third-party: online tracking</b></p> <p>Google also collects data about you from other third-party organisations i.e. other businesses or charities when that organisation has a relationship with Google.</p> <p>There are several ways in which this can happen:</p> <ul style="list-style-type: none"> <li>• Information about you is collected by these organisations (e.g. through their websites, apps, loyalty programs etc.) and is then shared directly with Google. They offer this information to Google because Google runs advertisements for these organisations through various Google services.</li> <li>• Information about your activity and interaction with these organisations is collected by Google using technologies such as cookies that these organisations use on their websites or apps or if you log into their websites or apps using your Google log-in.</li> <li>• Information about you is collected by Google when you interact with either a Google advertising service or with an organisation or app that uses a Google advertising service.</li> </ul> <p>Google uses this data to target you with adverts that Google thinks you will be interested in.</p>

**Q25. Third party online tracking**

How comfortable are you for **[Facebook/Google]** to collect third party online tracking data in order to inform targeted adverts? Please mark your answer on a scale of 1 to 10 where 1 is ‘Not comfortable at all’ and 10 is ‘Completely comfortable’

- 1 Not comfortable at all
- 2
- 3
- 4
- 5
- 6

- 7
- 8
- 9
- 10 Completely comfortable

Slide 3 Facebook

Slide 3 Google

Third-party customer lists

Another way that Facebook collects data about you to target its adverts is by matching your Facebook profile to customer lists provided by an organisation to Facebook.

You may be on an organisation's customer list if you signed up for an email newsletter, made a purchase at a retail shop or signed up for a voucher or discount.

A customer list tells Facebook that you have engaged with this organisation in some way, but Facebook gets no more information than this.

Organisations upload information such as your phone number or email address via a customer list and this can be matched to your Facebook profile. These lists are put into code so that your details can't be seen by anyone, but a computer programme can tell if the details match a Facebook profile.

Third-party customer lists

Another way that Google collects data about you to target its adverts is by matching your Google profile to customer lists provided by an organisation to Google.

You may be on an organisation's customer list if you signed up for an email newsletter, made a purchase at a retail shop or signed up for a voucher or discount.

A customer list tells Google that you have engaged with this organisation in some way, but Google gets no more information than this.

Organisations upload information such as your name, phone number or email address via a customer list and this can be matched to your Google profile. These lists are put into code so that your details can't be seen by anyone, but a computer programme can tell if the details match a Google profile.

Q26. Third party customer lists

How comfortable are you for [Facebook/Google] to collect third party customer list data in order to inform targeted adverts? Please mark your answer on a scale of 1 to 10 where 1 is 'Not comfortable at all' and 10 is 'Completely comfortable'

- 1 Not comfortable at all
- 2
- 3
- 4
- 5
- 6
- 7
- 8
- 9
- 10 Completely comfortable

Informed Choice

Q27. Suppose you were given the following choice.

IF OPTIONORDER=1

- In Option A, you would share data collected through all three ways with [Facebook/Google] and would see targeted adverts.
- In Option B, you would not share any of your data with [Facebook/Google] for the purposes of showing targeted adverts and would hence see generic adverts.

IF OPTIONORDER=2

- In Option A, you would not share any of your data with [Facebook/Google] for the purposes of showing targeted adverts and would hence see generic adverts.
- In Option B, you would share data collected through all three ways with [Facebook/Google] and would see targeted adverts.

IF REWARDFEE=REWARD Please note: there is no reward for choosing to see targeted adverts in this situation.

**IF REWARDFEE=FEE** Please note: there is no fee for choosing to see generic adverts in this situation.

Which option would you choose, A or B?

**IF OPTIONORDER=1, SHOW OPTION A="See targeted adverts"; OPTION B="See generic adverts"**

**IF OPTIONORDER=2, SHOW OPTION B="See targeted adverts"; OPTION A="See generic adverts"**

For Facebook

Type of adverts shown	Option A Full data sharing	Option B No data sharing
	Targeted	Generic
<i>Where information about you is collected from:</i> <b>1<sup>st</sup> party data collection:</b> Instagram/ Facebook activity i.e. which posts you like, content you watch or engage with, places you 'check-in' to and your profile data  <b>3<sup>rd</sup> party online tracking:</b> information from tracking you on other 3 <sup>rd</sup> party organisations' websites and apps  <b>3<sup>rd</sup> party customer lists:</b> matching your profile with customer lists uploaded by other organisations you have used/signed up to	YES	NO
	YES	NO
	YES	NO
	<input type="radio"/>	<input type="radio"/>

For Google

Type of adverts shown	Option A Full data sharing	Option B No data sharing
	Targeted	Generic
<i>Where information about you is collected from:</i> <b>1<sup>st</sup> party data collection:</b> Through Google services and Android mobile devices using Google's operating system  <b>3<sup>rd</sup> party online tracking:</b> information from tracking you on other 3 <sup>rd</sup> party organisations' websites and apps  <b>3<sup>rd</sup> party customer lists:</b> matching your profile with customer lists uploaded by other organisations you have used/signed up to	YES	NO
	YES	NO
	YES	NO
	<input type="radio"/>	<input type="radio"/>

**IF REWARDFEE=FEE SKIP TO Q34**

**Q29 [REWARDFEE=REWARD ONLY]** Now suppose you were given the following choice. This is the same as before except that in this case you would be offered a monthly reward for choosing to receive targeted adverts. As before, this may take the form of either:

- A payment through your PayPal account

- Amazon vouchers

Which option would you choose, A or B?

IF OPTIONORDER=1, SHOW OPTION A="Targeted"; OPTION B="Generic"

IF OPTIONORDER=2, SHOW OPTION B="Targeted"; OPTION A="Generic"

For Facebook

	Option A Full data sharing	Option B No data sharing
Type of adverts shown	Targeted	Generic
Where information about you is collected from:		
1 <sup>st</sup> party data collection: Instagram/ Facebook activity i.e. which posts you like, content you watch or engage with, places you 'check-in' to and your profile data	YES	NO
3 <sup>rd</sup> party online tracking: information from tracking you on other 3 <sup>rd</sup> party organisations' websites and apps	YES	NO
3 <sup>rd</sup> party customer lists: matching your profile with customer lists uploaded by other organisations you have used/signed up to	YES	NO
Value of reward paid to you each month ⓘ	BASECOST <sup>(1)</sup>	£0
	<input type="radio"/>	<input type="radio"/>

For Google

	Option A Full data sharing	Option B No data sharing
Type of adverts shown	Targeted	Generic
Where information about you is collected from:		
1 <sup>st</sup> party data collection: Through Google services and Android mobile devices using Google's operating system	YES	NO
3 <sup>rd</sup> party online tracking: information from tracking you on other 3 <sup>rd</sup> party organisations' websites and apps	YES	NO
3 <sup>rd</sup> party customer lists: matching your profile with customer lists uploaded by other organisations you have used/signed up to	YES	NO
Value of reward paid to you each month ⓘ	BASECOST	£0
	<input type="radio"/>	<input type="radio"/>

ⓘ HOVERBUTTON TEXT: If you change your mind you will continue to receive targeted adverts until the end of the month but can then opt out from targeted adverts. You will then receive no further rewards from then on.

Q31a How sure are you that this is the choice you would make in practice if this choice was really offered to you?

Very sure

Somewhat sure  
Not very sure  
Not at all sure

**BASECOST2\_INFORMED**  
=2\*BASECOST IF Q29="Generic"  
=0.5\*BASECOST IF Q29="Targeted"

---

Q32 Now suppose you were given the following choice. Which option would you choose, A or B?

[SHOW SAME CHOICE AS AT Q29 BUT WITH BASECOST REPLACED WITH BASECOST2\_INFORMED]

---

Q33 Now suppose you were given the following choice. This is the same as before except that in this case you would have multiple options for how much data you would be sharing with [Facebook/Google].

Which option would you choose?

**IF OPTIONORDER=1, SHOW:**

OPTION A="Full data sharing"  
OPTION B="1<sup>st</sup> party data and 3<sup>rd</sup> party customer lists"  
OPTION C="1<sup>st</sup> party data and 3<sup>rd</sup> party online tracking"  
OPTION D="1<sup>st</sup> party data only"  
OPTION E="No data sharing"

**IF OPTIONORDER=2, SHOW**

OPTION A="No data sharing"  
OPTION B="1<sup>st</sup> party data only"  
OPTION C="1<sup>st</sup> party data and 3<sup>rd</sup> party online tracking"  
OPTION D="1<sup>st</sup> party data and 3<sup>rd</sup> party customer lists"  
OPTION E="Full data sharing"

Facebook

	Option A Full data sharing	Option B 1 <sup>st</sup> party data and 3 <sup>rd</sup> party customer lists	Option C 1 <sup>st</sup> party data and 3 <sup>rd</sup> party online tracking	Option D 1 <sup>st</sup> party data only	Option E No data sharing
Type of adverts shown	Targeted	Targeted	Targeted	Targeted	Generic
Where information about you is collected from:					
1 <sup>st</sup> party data collection: Instagram/ Facebook activity i.e. which posts you like, content you watch or engage with, places you 'check-in' to and your profile data	YES	YES	YES	YES	NO
3 <sup>rd</sup> party online tracking: information from tracking you on other 3 <sup>rd</sup> party organisations' websites and apps	YES	NO	YES	NO	NO
3 <sup>rd</sup> party customer lists: matching your profile with customer lists uploaded by other organisations you have used/signed up to	YES	YES	NO	NO	NO
Value of reward paid to you each month ₤	BASECOST <sup>10</sup>	0.75*BASECOST	0.75*BASECOST	0.5*BASECOST	£0
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Google

	Option A Full data sharing	Option B 1 <sup>st</sup> party data and 3 <sup>rd</sup> party customer lists	Option C 1 <sup>st</sup> party data and 3 <sup>rd</sup> party online tracking	Option D 1 <sup>st</sup> party data only	Option E No data sharing
Type of adverts shown	Targeted	Targeted	Targeted	Targeted	Generic
Where information about you is collected from:					
1 <sup>st</sup> party data collection: Through Google services and Android mobile devices using Google's operating system	YES	YES	YES	YES	NO
3 <sup>rd</sup> party online tracking: information from tracking you on other 3 <sup>rd</sup> party organisations' websites and apps	YES	NO	YES	NO	NO
3 <sup>rd</sup> party customer lists: matching your profile with customer lists uploaded by other organisations you have used/signed up to	YES	YES	NO	NO	NO
Value of reward paid to you each month ₤	BASECOST <sup>10</sup>	0.75*BASECOST	0.75*BASECOST	0.5*BASECOST	£0
	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

SKIP TO 'About you'

Q34 [REWARD FEE = FEE ONLY] Now suppose you were given the following choice. This is the same as before except that in this case you would pay a monthly subscription fee for choosing to receive generic adverts rather than targeted adverts.

Which option would you choose, A or B?

IF OPTIONORDER=1, SHOW OPTION A="Targeted"; OPTION B="Generic"  
 IF OPTIONORDER=2, SHOW OPTION B="Targeted"; OPTION A="Generic"

For Facebook

	Option A Full data sharing	Option B No data sharing
Type of adverts shown	Targeted	Generic
Where information about you is collected from:		
1 <sup>st</sup> party data collection: Instagram/ Facebook activity i.e. which posts you like, content you watch or engage with, places you 'check-in' to and your profile data	YES	NO
3 <sup>rd</sup> party online tracking: information from tracking you on other 3 <sup>rd</sup> party organisations' websites and apps	YES	NO
3 <sup>rd</sup> party customer lists: matching your profile with customer lists uploaded by other organisations you have used/signed up to	YES	NO
Monthly subscription fee ⓘ	£0	BASECOST <sup>(1)</sup>
	<input type="radio"/>	<input type="radio"/>

For Google

	Option A Full data sharing	Option B No data sharing
Type of adverts shown	Targeted	Generic
Where information about you is collected from:		
1 <sup>st</sup> party data collection: Through Google services and Android mobile devices using Google's operating system	YES	NO
3 <sup>rd</sup> party online tracking: information from tracking you on other 3 <sup>rd</sup> party organisations' websites and apps	YES	NO
3 <sup>rd</sup> party customer lists: matching your profile with customer lists uploaded by other organisations you have used/signed up to	YES	NO
Monthly subscription fee ⓘ	£0	BASECOST <sup>(1)</sup>
	<input type="radio"/>	<input type="radio"/>

ⓘ HOVERBUTTON TEXT: If you change your mind you will continue to receive generic adverts until the end of the month but can then opt in to targeted adverts. You will then have no subscription fee to pay from then on.

Q36a How sure are you that this is the choice you would make in practice if this choice was really offered to you?

- Very sure
- Somewhat sure

Not very sure  
Not at all sure

DP: CREATE BASECOST2\_INFORMED  
=2\*BASECOST IF Q34="Generic"  
=0.5\*BASECOST IF Q34="Targeted"

---

Q37 Now suppose you were given the following choice. Which option would you choose, A or B?  
[SHOW SAME CHOICE AS AT Q34 BUT WITH BASECOST REPLACED WITH BASECOST2\_INFORMED]

---

Q38 Now suppose you were given the following choice. This is the same as before except that in this case you would have multiple options for where [Facebook/Google] can collect the data they use to target you with adverts

Which option would you choose?

IF OPTIONORDER=1, SHOW:  
OPTION A="Full data sharing"  
OPTION B="1<sup>st</sup> party data and 3<sup>rd</sup> party customer lists"  
OPTION C="1<sup>st</sup> party data and 3<sup>rd</sup> party online tracking"  
OPTION D="1<sup>st</sup> party data only"  
OPTION E="No data sharing"

IF OPTIONORDER=2, SHOW  
OPTION A="No data sharing"  
OPTION B="1<sup>st</sup> party data only"  
OPTION C="1<sup>st</sup> party data and 3<sup>rd</sup> party online tracking"  
OPTION D="1<sup>st</sup> party data and 3<sup>rd</sup> party customer lists"  
OPTION E="Full data sharing"

FACEBOOK

	Option A Full data sharing	Option B 1 <sup>st</sup> party data and 3 <sup>rd</sup> party customer lists	Option C 1 <sup>st</sup> party data and 3 <sup>rd</sup> party online tracking	Option D 1 <sup>st</sup> party data only	Option E No data sharing
Type of adverts shown	Targeted	Targeted	Targeted	Targeted	Generic
Where information about you is collected from:					
1 <sup>st</sup> party data collection: Instagram/Facebook activity i.e. which posts you like, content you watch or engage with, places you 'check-in' to and your profile data	YES	YES	YES	YES	NO
3 <sup>rd</sup> party online tracking: information from other 3 <sup>rd</sup> party organisations you use that have a relationship with Facebook	YES	NO	YES	NO	NO
3 <sup>rd</sup> party customer lists: matching your profile with customer lists uploaded by other organisations you have used/signed up to	YES	YES	NO	NO	NO
Monthly subscription fee	£0	0.25*BASECOST	0.25*BASECOST	0.5*BASECOST	BASECOST

○ ○ ○ ○ ○

GOOGLE

	Option A Full data sharing	Option B 1 <sup>st</sup> party data and 3 <sup>rd</sup> party customer lists	Option C 1 <sup>st</sup> party data and 3 <sup>rd</sup> party online tracking	Option D 1 <sup>st</sup> party data only	Option E No data sharing
Type of adverts shown	Targeted	Targeted	Targeted	Targeted	Generic
Where information about you is collected from:					
1 <sup>st</sup> party data collection: Through Google services and Android mobile devices using Google's operating system	YES	YES	YES	YES	NO
3 <sup>rd</sup> party online tracking: information from tracking you on other 3 <sup>rd</sup> party organisations' websites and apps	YES	NO	YES	NO	NO
3 <sup>rd</sup> party customer lists: matching your profile with customer lists uploaded by other organisations you have used/signed up to	YES	YES	NO	NO	NO
Monthly subscription fee	£0	0.25*BASECOST	0.25*BASECOST	0.5*BASECOST	BASECOST

○ ○ ○ ○ ○

About you

Thank you for helping us with this questionnaire. We would now like to ask a few questions about yourself, which will only be used for analysis purposes and, like anything else you told us today, will not be shared with anyone not associated with the research project.

Q39 Thinking about the devices that you use [LIST OF DEVICES PARTICIPANT USES FROM Q8a] do you have ad blockers set up on any of these devices?

- Yes
- No
- Don't know

---

Q39a Were you aware that in April 2021 Apple brought in new privacy changes which gives users the choice of whether the apps on their phone can track their activity across other company's apps and websites?

- Yes I was aware of this
- No I was not aware of this

---

Q40 What is your ethnic group?

- White English/Welsh/Scottish/Northern Irish/British
- White Irish
- White Gypsy or Irish Traveller
- Any other White background, please describe
- White and Black Caribbean
- White and Black African
- Any other Mixed/Multiple ethnic background, please describe
- Indian
- Pakistani
- Bangladeshi
- Chinese
- Any other Asian background, please describe
- African
- Caribbean
- Any other Black/African/Caribbean background, please describe
- Arab
- Any other ethnic group, please describe
- Prefer not to answer

---

Q41 What is your household annual income?

- Less than £20,000
- Between £20,000 and £39,999
- Between £40,000 and £59,999
- Between £60,000 and £79,999
- Between £80,000 and £99,999
- More than £100,000
- Prefer not to say

---

Q42 What is your highest educational qualification?

- Doctorate
- Postgraduate degree
- Undergraduate degree
- Higher education below a degree
- GCE A-Level / AS-Level or equivalent
- GCSE grade A\*-C / GCE O-Level or equivalent
- GCSE grade D-G / CSE or equivalent
- Other
- No formal qualifications
- Prefer not to say

---

Q43 Which of these, if any, impact or limit your daily activities or the work that you can do?

Hearing? Poor hearing, partial hearing or are deaf

Eyesight? Poor vision, colour blindness, partial sight or are blind

Dexterity? Limited ability to reach/ difficulty opening things with your hands/ difficulty using a telephone handset/ television remote/ computer keyboard etc

Breathing? Breathlessness or chest pains

Mental abilities? Such as learning, understanding, concentration, memory, communicating, cognitive loss or deterioration.

Social/ behavioral? Conditions associated with this such as autism, attention deficient disorder

Your mental health? Anxiety, depression, or trauma related conditions

Other illness/ conditions that impact your daily activities or the work you can do

Nothing – no impairments or conditions that impact your daily life or the work you can do

Prefer not to say

Don't know

## CLOSING PAGE

Thank you for completing the survey. This research was conducted under the terms of the MRS code of conduct and is completely confidential.

Please press the forward button at the bottom of the page to submit your data and exit the survey.

## Appendix B – Sample Demographics

The four tables below show a comparison between the population and sample statistics for gender, age, region participants currently live in and the social grade.<sup>34</sup>

**Table 18:** Q2. What gender do you identify as?

Gender	Population: UK	Sample		
		All (Base: 4,003)	Facebook (Base: 1,996)	Google (Base: 2,007)
Female	51%	54%	55%	53%
Male	49%	46%	45%	47%
<b>Total</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>

Note: 10 participants who chose 'Other' and 1 participant who chose 'Prefer not to say' are not included in the sample statistics

**Table 19:** Q3. How old are you?

Age group	Population: UK	Sample		
		All (Base: 4,013)	Facebook (Base: 2,005)	Google (Base: 2,008)
18-29	20.6%	14.5%	14.4%	14.6%
30-44	26.0%	25.2%	30.3%	20.2%
45-64	32.5%	35.8%	30.3%	41.3%
65+	20.9%	24.5%	25.1%	23.9%
<b>Total</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>

Note: 1 participant who chose 'Prefer not to say' is not included in the sample statistics

34 The population statistics are extracted from the NOMIS website at <https://www.nomisweb.co.uk/census/2011>

**Table 20:** Q4. Please select from the list below the region you are currently living in

Region	Population: UK	Sample		
		All (Base: 4,014)	Facebook (Base: 2,005)	Google (Base: 2,009)
North East	4.1%	4.3%	3.6%	5.1%
North West	11.2%	11.1%	12.1%	10.2%
Yorkshire and The Humber	8.4%	7.8%	7.9%	7.7%
East Midlands	7.2%	7.6%	7.3%	8.0%
West Midlands	8.9%	9.1%	9.0%	9.2%
East of England	9.3%	8.7%	8.9%	8.5%
London	12.9%	12.5%	11.8%	13.3%
South East	13.7%	14.7%	14.3%	15.2%
South West	8.4%	9.1%	9.7%	8.5%
Wales	4.8%	5.1%	5.6%	4.5%
Scotland	8.4%	8.0%	8.3%	7.8%
Northern Ireland	2.9%	1.8%	1.5%	2.1%
<b>Total</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>

**Table 21:** Q7b.Socio-economic grades

Social grade	Population: UK	Sample		
		All (Base: 3,000)	Facebook (Base: 1,485)	Google (Base: 1,515)
AB	22.2%	21.5%	18.4%	24.5%
C1	30.8%	30.3%	27.9%	32.5%
C2	20.9%	21.1%	23.8%	18.5%
DE	26.1%	27.1%	29.8%	24.5%
<b>Total</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>	<b>100.0%</b>

Note: The population numbers are for all UK residents aged 16–64 years while the sample numbers pertain to all participants aged 18–64 years and exclude 30 participants aged 18–64 for whom social grade could not be coded because they chose not to answer one of the social grade questions

A comparison of the sample and population statistics for age, gender, region and social grade indicates a reasonably close match between Google and Facebook users, as represented in the sample, and the full population of the UK, as represented by national statistics. The proportion of participants aged 18–29 was somewhat lower than the corresponding population proportion.

Participants were randomly assigned to take the survey as Facebook or Google users unless they used Facebook (Google) services less than once a month or never, in which case they were assigned to the Google (Facebook) group. Facebook and Google users were similar in terms of gender and region. Facebook users tended to be younger and of lower social grade compared to Google users.

# Appendix C – Econometric Modelling of Contingent Valuation Choices and CRR Valuation

## Introduction

The core results from the research have been obtained via an econometric analysis of responses to the contingent valuation choice questions. This appendix provides details of the econometric model specifications adopted and the estimation results.

## Choice Model Specification

In the binary contingent valuation questions, each participant made 6 choices between receiving ‘Generic’ and ‘Targeted’ adverts (3 questions uninformed and 3 questions informed). The choices were analysed via econometric discrete choice models, with choice as the dependent variable, a {1,0} variable indicating whether ‘Generic’ adverts were chosen in any given question. Choices are interpreted as indicating that the ‘utility’ of the chosen option (‘Generic’ or ‘Targeted’) is greater than the ‘utility’ of the other options. This interpretation follows the principles of random utility theory (see Train 2003).

The utility  $U_{ijk}$  that individual  $i$  obtains in choice situation  $j$  from option  $k \in \{G, T\}$  (one of ‘Generic’ or ‘Targeted’) is assumed to be a function of a preference parameter  $\beta_k$  (a constant), the ‘marginal utility of income’  $\beta_{Cost}$ , income  $M$ , the amount of the fee/reward, and a random component  $e_{ijk}$ :

$$U_{ijG} = \beta_G + \beta_{Cost}(M - Fee_{ij}) + e_{ijG} \quad (2)$$

$$U_{ijT} = \beta_T + \beta_{Cost}M + e_{ijT} \quad (3)$$

‘Generic’ is chosen if it yields a greater utility than ‘Targeted’. Therefore, the probability that ‘Generic’ is chosen is given by

$$P(U_{ijG} - U_{ijT} > 0) = P(\beta_G - \beta_T - \beta_{Cost}Fee_{ij} > e_{ijT} - e_{ijG}) \quad (4)$$

Because only differences in utility matter, as shown in equation (4), the choice model can be written as

$$U_{ijG} = \beta_G - \beta_T - \beta_{Cost}Fee_{ij} + e_{ijG} \quad (5)$$

$$U_{ijT} = e_{ijT} \quad (6)$$

Analogous formulae for the ‘reward’ case imply that the model can be written as:

$$U_{ijG} = \beta_0 - \beta_{Cost}Cost_{ij} + e_{ijG} \quad (7)$$

$$U_{ijT} = e_{ijT} \quad (8)$$

The choice probability for ‘Generic’ is given by

$$P(U_{ijG} - U_{ijT} > 0) = P(\beta_0 - \beta_{Cost}Cost_{ij} > e_{ijT} - e_{ijG}) \quad (9)$$

where  $\beta_0 \equiv \beta_G - \beta_T$  is an intercept term—we are only able to estimate the difference—and the cost is that of ‘Generic’ relative to ‘Targeted’, i.e., the amount of the fee or reward. We estimate the parameters  $\beta_0$  and  $\beta_{Cost}$  using logit models which are obtained assuming  $e_{ijG}$  and  $e_{ijT}$  follow independent Gumbel distributions.

### CRR Valuation Approach

Our approach to valuing the CRR is based on the premise that the CRR will expand the ‘choice set’ of Facebook and Google users by effectively giving them the choice not to share their data for targeted advertising and to receive generic adverts only. Hence, our estimate is given by the change in expected consumer surplus arising from a change in the choice set from  $\{T\}$  to  $\{G, T\}$ , where the ‘Generic’ option would be available at zero cost.

Applying the log-sum formula for the change in expected consumer surplus (e.g., Train 2003)

$$\Delta E(CS_i) = \frac{1}{\alpha_i} \left[ \ln \left( \sum_{j=1}^J e^{V_{ij}} \right) - \ln \left( \sum_{k=1}^K e^{V_{ik}} \right) \right] \quad (10)$$

where  $\alpha_i$  is the marginal utility of money, and  $V_{ij}$  and  $V_{ik}$  denote the deterministic components of the utility of the options available for choice in the post-CRR and pre-CRR scenarios, i.e.,

$$V_{ijG} = \beta_0 \quad (11)$$

$$V_{ijT} = 0 \quad (12)$$

yields

$$\Delta E(CS_i) = \frac{1}{\beta_{Cost}} [\ln(1 + e^{\beta_0})] \quad (13)$$

Our estimate is given by

$$\Delta \widehat{E}(CS_i) = \frac{1}{\widehat{\beta}_{Cost}} [\ln(1 + e^{\widehat{\beta}_0})] \quad (14)$$

where  $\widehat{\beta}_0$  and  $\widehat{\beta}_{Cost}$  are coefficient estimates from a logit model<sup>35</sup>.

In the random-effects (RE) logit model the intercept is allowed to vary across individuals, and it is given by  $\beta_0 + v_i$ , where  $v_i \sim N(0, \sigma_v^2)$ . That is, the intercept is given by the sum of a constant term and a zero-mean normally distributed random effect allowing

<sup>35</sup> Some of our models include ‘interaction’ terms allowing  $\widehat{\beta}_0$  and  $\widehat{\beta}_{Cost}$  to vary depending on (a) whether choices were informed or uninformed; (b) whether participants were sure that they would make the same choices in practice; (c) whether participants had reviewed their data settings.

preferences regarding 'Generic' and 'Targeted' adverts to differ across individuals. The change in expected consumer surplus is given by

$$\Delta E(CS_i) = \frac{1}{\beta_{Cost}} [\ln(1 + e^{\beta_0 + v_i})] \quad (15)$$

Let

$$\begin{aligned} \mu &\equiv E \left[ \frac{1}{\beta_{Cost}} [\ln(1 + e^{\beta_0 + v_i})] \right] \\ &= \int_{-\infty}^{\infty} \frac{1}{\beta_{Cost}} [\ln(1 + e^{\beta_0 + v_i})] \frac{e^{-v_i^2/2\sigma_v^2}}{\sqrt{2\pi}\sigma_v} dv_i \end{aligned} \quad (16)$$

denote the mean of the change in expected consumer surplus, i.e., the mean value of the CRR. We estimate  $\mu$  via simulation by taking 2,000 independent draws from  $N(0, \hat{\sigma}_v^2)$ , applying the transformation given in equation (15) using the parameter estimates  $\hat{\beta}_0, \hat{\beta}_{Cost}, \hat{\sigma}_v$  from the RE logit model, and calculating the mean of the transformed draws.

For statistical inference involving the mean change in expected consumer surplus derived from our preferred RE logit models we employ the Krinsky-Robb approach (see, e.g., Haab and McConnell 2002), carrying out the following steps:

- a) estimating the relevant RE logit model;
- b) drawing 1,000 parameter vectors  $(\tilde{\beta}_0, \tilde{\beta}_C, \tilde{\sigma}_v^2)$  from the (estimated) distribution of the estimators;
- c) simulating  $\tilde{\mu}$  conditional on each parameter vector, as explained above.

The standard deviation of the distribution of simulated  $\tilde{\mu}$ s is used to construct confidence intervals and hypothesis tests.

### Econometric Model Building

Separate models were estimated for participants who were asked to pay a fee to avoid receiving targeted adverts and participants who were offered a reward for receiving targeted adverts, as choice patterns were substantially different between fee and reward treatments. The models are shown in Table 22 and Table 23 for fee and reward treatments, respectively, and differ in terms of estimation samples, independent variables, and assumptions around the intercept (fixed vs random). The estimated (mean) change in expected consumer surplus is shown in the bottom rows.

The models in columns (1) to (8) use only responses to the first contingent valuation question which included a fee/reward, i.e., the second question (Q2) in each set of uninformed and informed choices. (Q1 did not involve any payments.) We estimate standard logit models for informed and uninformed choices in columns (1) and (2), respectively, and models combining both choices in columns (3) and (4), allowing for differences in the preference parameters between uninformed and informed scenarios. The cost coefficient has the expected negative sign, but it is not statistically significant in two models for the fee sample. The positive and statistically significant coefficient on

the 'Informed choice' dummy variable indicates a stronger preference for privacy once participants are informed about the extent of data collection by Facebook/Google.

The estimated welfare impact of the CRR, calculated applying equation (14), ranges between £1.47 per month and £4.73 per month depending on payment vehicle and information about data collection.

The standard logit model assumes that the preference parameters are 'fixed', i.e., do not vary across individuals. Consider the intercept  $\beta_0$  which captures the deterministic component of utility when the cost of choosing generic adverts is zero, as is the case in the first contingent valuation question (Q1), which does not involve any payments, and hence represents a benchmark scenario. A positive (negative) intercept implies a greater tendency to choose 'Generic' ('Targeted') in that scenario.

There are good reasons to expect the sign of  $\beta_0$  to vary across individuals. Some users may be relaxed about the collection of their data, while others may be concerned. Some users will find personalised adverts more 'relevant' and helpful in reducing 'search costs', while others will find these adverts to be intrusive and 'repetitive'. Some users may regard generic adverts as being more varied and interesting, while others will see generic adverts as being irrelevant and useless. The fact that a substantial proportion of participants chose 'Targeted' even when there was no monetary incentive for doing so (see Table 4), confirms the view that preferences about digital advertising and the use of personal data, specifically, as measured by the sign of the intercept term, vary across users.

The random effects (RE) logit model in columns (5) to (12) overcomes two limitations of the standard logit model<sup>36</sup>. First, the intercept  $\beta_0$  becomes a 'random effect' which varies across individuals according to a normal distribution allowing for heterogeneous preferences. Second, a sequence of choices can be modelled meaning we are able to use the full set of 3 contingent valuation responses (Q1, Q2, and Q3, of which only the latter two included a fee/reward) for each of the uninformed and informed scenarios, i.e., 6 choices per participant, overall.

As an intermediate step, in columns (5) to (8) we use the same data as for the standard logit models in columns (1) to (4). Unsurprisingly, models (5) and (6) yield very similar estimates as models (1) and (2). This is because the RE model cannot exploit the panel nature of the data, as it is estimated on one choice per participant.

For the random-effects models in columns (5) to (12) we need to consider that the change in expected consumer surplus given by the log-sum formula is no longer constant across individuals, as shown in equation (15). Due to the inclusion of a random-effect term in the log-sum expression, the change in expected consumer surplus varies across the population according to an asymmetric distribution. We calculate the mean of the distribution via simulation as explained above under 'CRR Valuation Approach'.

---

<sup>36</sup> The model is analogous to the random effects probit, which was proposed for contingent valuation analysis by Alberini et al. (1997). See also Haab and McConnell (2002).

The models in columns (7) and (8), which are estimated on the combined sample of Q2 uninformed and informed choices give a better fit and produce lower estimates of the mean increase in consumer surplus due to the CRR, except in the case of the consumer surplus under informed choices in the fee sample (see column (8)).

The models estimated on the full set of responses (i.e., Q1, Q2, and Q3 for both uninformed and informed choice) are shown in columns (9) to (12). We include a Q3 dummy variable, which marks the contingent valuation follow-up question (i.e., the second payment question), as suggested by Alberini et al. (1997) to allow for a shift in WTP between the first payment question and the follow-up question. Model (11) also include a full set of interaction terms with an 'Informed' dummy allowing for information about data collection to affect preferences.

The coefficient on the cost variables is negative and highly statistically significant in columns (9) to (12) indicating that participants' choices were sensitive to the cost factor. The coefficient on the 'Informed' dummy is positive and statistically significant in columns (11) and (12) indicating, as expected, a greater preference for 'Generic' in informed choices compared to uninformed choices. The coefficient on the Q3 dummy is negative and statistically significant in the fee sample, in line with typical findings, and positive and only marginally significant in one model only for the reward sample.

Models (9) to (12) are based on all binary contingent valuation choices between 'Generic' and 'Targeted' made by participants in the survey. The models allow preferences regarding data collection/type of adverts to vary across individuals and can accommodate changes in WTP for data privacy/generic adverts as may occur between first and follow-up payment questions in double-bounded contingent valuation exercises. Models (11) and (12), which are estimated on the combined sample of uninformed and informed choices, have a better fit, as measured by the log-likelihood, than models estimated on sub-samples. We select model (11) as our preferred model on the basis that the interaction between 'Cost' and 'Informed choice' is statistically significant in the reward sample, while it does not affect our welfare estimates in the fee sample.

The mean change in expected consumer surplus due to the CRR based on our preferred (baseline) model is £1.12 (uninformed) and £1.86 (informed) per user per month for the fee sample and £2.26 (uninformed) and £5.77 (informed) per user per month for the reward sample.

**Table 22:** Econometric models of choice between ‘Generic’ and ‘Targeted’ adverts: Fee sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variable	Logit Q2 Uninformed	Logit Q2 Informed	Logit Q2 Uninformed & Informed	Logit Q2 Uninformed & Informed	RE logit Q2 Uninformed	RE logit Q2 Informed	RE logit Q2 Uninformed & Informed	RE logit Q2 Uninformed & Informed	RE logit Q1&Q2&Q3 Uninformed	RE logit Q1&Q2&Q3 Informed	RE logit Q1&Q2&Q3 Uninformed & Informed	RE logit Q1&Q2&Q3 Uninformed & Informed
Informed choice			1.042*** (0.134)	0.911*** (0.079)			1.778*** (0.229)	1.514*** (0.144)			0.849*** (0.060)	0.861*** (0.049)
Cost of ‘Generic’	-0.073 (0.052)	-0.136*** (0.039)	-0.073 (0.052)	-0.115*** (0.036)	-0.073 (0.052)	-0.137*** (0.039)	-0.110 (0.079)	-0.192*** (0.060)	-0.532*** (0.062)	-0.569*** (0.053)	-0.633*** (0.068)	-0.624*** (0.050)
Cost x Informed			-0.064 (0.052)				-0.124 (0.083)				0.015 (0.060)	
Q3 dummy									-1.475*** (0.072)	-1.553*** (0.063)	-1.521*** (0.076)	-1.524*** (0.080)
Q3 dummy x Informed											-0.022 (0.082)	-0.016 (0.087)
Constant	-2.186*** (0.134)	-1.144*** (0.097)	-2.186*** (0.134)	-2.100*** (0.107)	-2.230*** (0.136)	-1.151*** (0.098)	-3.942*** (0.283)	-3.759*** (0.251)	-0.348*** (0.054)	0.480*** (0.059)	-0.315*** (0.059)	-0.322*** (0.050)
ln( $\sigma_v^2$ )					-2.198 (0.000)	-3.629 (0.000)	1.849*** (0.155)	1.840*** (0.155)	-0.219 (0.227)	0.247 (0.166)	0.407*** (0.110)	0.407*** (0.110)
Observations	2,000	2,000	4,000	4,000	2,000	2,000	4,000	4,000	6,000	6,000	12,000	12,000
Participants					2,000	2,000	2,000	2,000	2,000	2,000	2,000	2,000
Log likelihood	-592.4	-971.5	-1564	-1564	-592.4	-971.5	-1460	-1461	-2984	-3403	-6201	-6201
Pseudo R-squared	0.00171	0.00692	0.0335	0.0332	0.00171	0.00692	0.0977	0.0970	0.116	0.140	0.166	0.166
<b>(Mean) <math>\Delta E(CS)</math> (£/user/month)</b>												
Uninformed	£1.47	-	£1.47	£1.01	£1.46	-	£1.31	£0.84	£1.17	-	£1.12	£1.13
Informed	-	£2.03	£2.03	£2.32	-	£2.03	£1.85	£2.15	-	£1.93	£1.86	£1.85

Dependent variable: choice. Standard errors clustered by participant in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.  $\Sigma_v$  denotes the standard deviation of the random effects. Mean change in expected consumer surplus for random-effects (RE) logit models (5) to (12) estimated via simulation of expression (16), based on 2,000 draws from setting ‘Q3 dummy’ = 0. Variables: – Choice (of ‘Generic’) = 1 if ‘Generic’ (no data sharing) chosen; 0 otherwise  
 – Informed (choice) = 1 for Q1, Q2, Q3 of the set of informed choices; = 0 for Q1, Q2, Q3 of uninformed choices  
 – Cost of ‘Generic’ = £0 in Q1; = {£0.50, £1.00, £2.00, £3.00, £5.00} per month in Q2 and amount doubled or halved in Q3 depending on response to Q2; Cost of ‘Generic’ (no data sharing) relative to ‘Targeted’ (full data sharing)

**Table 23:** Econometric models of choice between ‘Generic’ and ‘Targeted’ adverts: Reward sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Variable	Logit Q2 Uninformed	Logit Q2 Informed	Logit Q2 Uninformed & Informed	Logit Q2 Uninformed & Informed	RE logit Q2 Uninformed	RE logit Q2 Informed	RE logit Q2 Uninformed & Informed	RE logit Q2 Uninformed & Informed	RE logit Q1&Q2&Q3 Uninformed	RE logit Q1&Q2&Q3 Informed	RE logit Q1&Q2&Q3 Uninformed & Informed	RE logit Q1&Q2&Q3 Uninformed & Informed
Informed choice			0.956*** (0.083)	1.025*** (0.049)			2.580*** (0.239)	2.639*** (0.178)			1.091*** (0.072)	1.301*** (0.064)
Cost of ‘Generic’	-0.199*** (0.035)	-0.166*** (0.028)	-0.199*** (0.035)	-0.179*** (0.027)	-0.200*** (0.036)	-0.167*** (0.028)	-0.467*** (0.088)	-0.451*** (0.072)	-0.481*** (0.044)	-0.502*** (0.042)	-0.512*** (0.039)	-0.374*** (0.025)
Cost x Informed			0.033 (0.033)				0.027 (0.079)				0.205*** (0.035)	
Q3 dummy									0.104 (0.068)	0.154* (0.085)	0.098 (0.071)	-0.008 (0.072)
Q3 dummy x Informed											-0.046 (0.082)	0.162** (0.082)
Constant	-0.775*** (0.090)	0.180** (0.078)	-0.775*** (0.090)	-0.816*** (0.078)	-0.780*** (0.090)	0.181** (0.078)	-2.097*** (0.243)	-2.132*** (0.225)	-0.106 (0.070)	1.439*** (0.123)	-0.111 (0.074)	-0.257*** (0.070)
ln( $\sigma_v^2$ )					-3.708 (0.000)	-3.922 (0.000)	2.694*** (0.136)	2.695*** (0.136)	1.602*** (0.097)	2.530*** (0.101)	1.802*** (0.066)	1.755*** (0.064)
Observations	2,014	2,014	4,028	4,028	2,014	2,014	4,028	4,028	6,042	6,042	12,084	12,084
Participants					2,014	2,014	2,014	2,014	2,014	2,014	2,014	2,014
Log likelihood	-1069	-1368	-2437	-2437	-1069	-1368	-2142	-2142	-3651	-3447	-6663	-6689
Pseudo R-squared	0.0157	0.0129	0.0562	0.0561	0.0157	0.0129	0.170	0.170	0.0989	0.167	0.203	0.200
<b>(Mean) <math>\Delta E(CS)</math> (£/user/month)</b>												
Uninformed	£1.90	-	£1.90	£2.04	£1.90	-	£1.79	£1.83	£2.25	-	£2.26	£2.85
Informed	-	£4.73	£4.73	£4.48	-	£4.73	£4.42	£4.34	-	£4.81	£5.77	£4.79

Dependent variable: *choice*. Standard errors clustered by participant in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.  $\Sigma_v$  denotes the standard deviation of the random effects. Mean change in expected consumer surplus for random-effects (RE) logit models (5) to (12) estimated via simulation of expression (16), based on 2,000 draws from setting ‘Q3 dummy’ = 0. Variables: - Choice (of ‘Generic’) = 1 if ‘Generic’ (no data sharing) chosen; 0 otherwise  
 - Informed (choice) = 1 for Q1, Q2, Q3 of the set of informed choices; = 0 for Q1, Q2, Q3 of uninformed choices  
 Cost of ‘Generic’ = £0 in Q1; = {£0.50, £1.00, £2.00, £3.00, £5.00} per month in Q2 and amount doubled or halved in Q3 depending on response to Q2; Cost of ‘Generic’ (no data sharing) relative to ‘Targeted’ (full data sharing)

## Sensitivity Tests and Final Models

Our preferred baseline models are random-effects logit models estimated on the full set of responses, i.e., Q1, Q2, and Q3 covering both uninformed and informed choices. The models are shown in column (11) in Table 22 and Table 23. Here, we carry out sensitivity tests on these models to arrive at our final models on which we base our estimates of CRR valuations. These tests and our final models are shown in Table 24 and Table 26 for the fee and reward samples, respectively.

Column (1) in Table 24 (fee sample) and Table 26 (reward sample) shows our preferred baseline model.

In column (2) we exclude participants whose questionnaire completion time was below the 10<sup>th</sup> percentile of completion times which was 3 minutes and 25 seconds. There is a concern that ‘speeders’ do not consider the options carefully enough or may even respond randomly to complete quickly.

In column (3) we look at how choices made by participants who were not sure they would make the same choices in a real choice situation impact our estimates of the change in consumer surplus due to implementation of a CRR. Several studies have found an inverse correlation between stated levels of confidence and ‘hypothetical bias’, i.e., the tendency for individuals to have a greater willingness to pay in choice situations which involve hypothetical payments as opposed to actual payments. We include a set of interaction terms with a dummy variable indicating low confidence<sup>37</sup>, allowing for different valuations of the CRR by those who were/were not sure of their choices. The CRR valuation of those who were ‘Very sure’ or ‘Somewhat sure’ is not substantially different from the overall valuation derived from the model in column (1) which does not include the level-of-confidence interaction terms.

In column (4) we combine the exclusion of ‘speeders’ and the choice confidence interaction terms, and we additionally include a set of interaction terms with a dummy variable indicating whether participants had reviewed their data settings, regardless of whether they also changed the settings<sup>38</sup>.

However, for these individuals, the CRR cannot be seen as providing an ‘expanded choice set’. For our calculations of the mean change in expected consumer surplus in column (5) we assume that

$$\Delta E(CS_i) = 0 \tag{17}$$

for individuals who have already reviewed their data settings. This is a conservative assumption, as those who have reviewed their settings may still experience a welfare

<sup>37</sup> The dummy variable takes the value of 1 for both the second (Q2) and third (Q3) contingent valuation question—Q1 does not involve any payments—if a participant stated in a follow-up question to Q2 that she/he was ‘Not very sure’ or ‘Not at all sure’ that she/he would make the same choice in practice. The confidence follow-up question was asked once for each set of informed and uninformed choices and, so, the ‘confidence dummy variable’ may take on different values for any given participant.

<sup>38</sup> This does not include those who reviewed their settings but were not sure how to change them.

gain from a more user-friendly way of setting data sharing preferences following implementation of the CRR.

These assumptions imply that the change in expected consumer surplus is a mixture of a continuous distribution and a degenerate distribution having mass 1 at 0. The mixing probabilities are given by the proportions who have reviewed/have not reviewed their data settings.

$$\begin{aligned} & \text{Reviewed} \sim \text{Bernoulli}(\theta) \\ & (\Delta E(CS_i) | \text{Reviewed} = 1) = 0 \\ & (\Delta E(CS_i) | \text{Reviewed} = 0) = \frac{1}{\beta_{Cost}} [\ln(1 + e^{\beta_0 + v_i})], \quad v_i \sim N(0, \sigma_v^2) \end{aligned} \tag{18}$$

Let the following expression again denote the mean of the change in expected consumer surplus, i.e., the mean value of the CRR, for those who did not review their data settings.

$$\begin{aligned} \mu & \equiv E \left[ \frac{1}{\beta_{Cost}} [\ln(1 + e^{\beta_0 + v_i})] \right] \\ & = \int_{-\infty}^{\infty} \frac{1}{\beta_{Cost}} [\ln(1 + e^{\beta_0 + v_i})] \frac{e^{-v_i^2/2\sigma_v^2}}{\sqrt{2\pi}\sigma_v} dv_i \end{aligned} \tag{19}$$

As explained above, we estimate  $\mu$  via simulation by taking 2,000 independent draws from  $N(0, \hat{\sigma}_v^2)$ , applying the transformation given in equation (15) based on the parameter estimates from the RE logit model, and calculating the mean of the transformed draws<sup>39</sup>.

Letting  $\hat{\theta}$  denote the sample proportion of those who reviewed their data settings, excluding those who were not sure how to change the settings, our estimate of the mean value of the CRR, is given by

$$E[\widehat{\Delta E(CS_i)}] = (1 - \hat{\theta}) \cdot \hat{\mu} \tag{20}$$

For inference about the mean change in expected consumer surplus we apply the Krinsky-Robb approach implementing the steps described above under ‘CRR Valuation Approach’ and, additionally, drawing bootstrap samples from our dataset to allow for sampling error in the estimated proportion  $\hat{\theta}$ . Thus, our procedure requires:

- a) estimating the relevant RE logit model;
- b) drawing 1,000 parameter vectors  $(\tilde{\beta}_0, \tilde{\beta}_C, \tilde{\sigma}_v^2)$  from the (estimated) distribution of the estimators;
- c) drawing 1,000 bootstrap samples from our dataset each yielding an estimate  $\tilde{\theta}$  of the proportion of those who reviewed their data settings;
- d) simulating  $\tilde{\mu}$  conditional on each parameter vector draw, as explained above;
- e) calculating  $E[\widehat{\Delta E(CS_i)}] = (1 - \tilde{\theta}) \cdot \tilde{\mu}$ .

<sup>39</sup> For informed valuations we use  $\beta_0 + \beta_{informed}$  and  $\beta_{Cost} + \beta_{Cost \times informed}$  in lieu of  $\beta_0$  and  $\beta_{Cost}$ , respectively.

The standard deviation of the simulated distribution of mean change in expected consumer surplus is used to construct confidence intervals and hypothesis tests.

Table 24: Sensitivity analysis and final model: Fee sample

	(1)	(2)	(3)	(4)	(5)
Variable	Baseline	Baseline & 'Speeders' (fastest 10%) excluded	Baseline & 'Not sure' interactions	(1)&(2)&(3) & 'Settings reviewed' interactions	(4) & 'Settings reviewed' implies CRR value of 0
Informed choice	0.849*** (0.060)	0.907*** (0.065)	0.899*** (0.064)	1.094*** (0.083)	1.094*** (0.083)
Cost of 'Generic'	-0.633*** (0.068)	-0.730*** (0.081)	-0.657*** (0.075)	-1.251*** (0.122)	-1.251*** (0.122)
Cost x Informed	0.015 (0.060)	0.055 (0.070)	0.002 (0.069)	0.248** (0.116)	0.248** (0.116)
Q3 dummy	-1.521*** (0.076)	-1.589*** (0.084)	-1.592*** (0.085)	-1.930*** (0.135)	-1.930*** (0.135)
Q3 dummy x Informed	-0.022 (0.082)	-0.027 (0.090)	-0.127 (0.093)	-0.194 (0.149)	-0.194 (0.149)
Not sure would make same choice in practice			-3.495*** (0.306)	-3.675*** (0.491)	-3.675*** (0.491)
Informed x Not sure			1.139*** (0.300)	1.070** (0.439)	1.070** (0.439)
Cost x Not sure			0.793*** (0.118)	1.078*** (0.214)	1.078*** (0.214)
Cost x Informed x Not sure			-0.240** (0.116)	-0.344* (0.188)	-0.344* (0.188)
Q3 dummy x Not sure			1.989*** (0.266)	2.029*** (0.368)	2.029*** (0.368)
Q3 dummy x Informed x Not sure			-0.478* (0.290)	-0.236 (0.379)	-0.236 (0.379)
Settings reviewed				-0.367*** (0.137)	-0.367*** (0.137)
Informed x Reviewed				-0.380*** (0.139)	-0.380*** (0.139)
Cost x Reviewed				0.902*** (0.143)	0.902*** (0.143)
Cost x Informed x Reviewed				-0.273** (0.132)	-0.273** (0.132)
Q3 dummy x Reviewed				0.734*** (0.181)	0.734*** (0.181)
Q3 dummy x Informed x Reviewed				0.002 (0.201)	0.002 (0.201)
Constant	-0.315*** (0.059)	-0.276*** (0.064)	-0.195*** (0.061)	0.007 (0.076)	0.007 (0.076)
$\ln(\sigma_v^2)$	0.407*** (0.110)	0.508*** (0.118)	0.554*** (0.106)	0.676*** (0.106)	0.676*** (0.106)
Observations	12,000	10,584	12,000	10,584	10,584
Participants	2,000	1,764	2,000	1,764	1,764
Log likelihood	-6201	-5369	-6035	-5050	-5050
Pseudo R-squared	0.166	0.181	0.188	0.230	0.230
<b>Mean <math>\Delta E(CS)</math> (£/user/month)</b>					
Uninformed	£1.12	£1.01	£1.19	£0.72	£0.50

**101 VALUE OF THE CHOICE REQUIREMENT REMEDY**

Uninformed: 95% CI	[£0.92-£1.31]	[£0.83-£1.20]	[£0.96-£1.43]	[£0.60-£0.85]	[£0.41-£0.59]
Informed	£1.86	£1.82	£1.95	£1.56	£1.09
Informed: 95% CI	[£1.62-£2.10]	[£1.58-£2.05]	[£1.66-£2.24]	[£1.36-£1.75]	[£0.95-£1.23]

Dependent variable: *choice*. Standard errors clustered by participant in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.  $\sigma_v$  denotes the standard deviation of the random effects. Mean change in expected consumer surplus calculated setting 'Not sure' = 0, 'Reviewed' = 0, 'Q3 dummy' = 0.

**Table 25:** Descriptive statistics – Fee sample

Variable		(1)	(2)	(3)	(4)	(5)
		Baseline	Baseline & 'Speeders' (fastest 10%) excluded	Baseline & 'Not sure' interactions	(1)&(2)&(3) & 'Settings reviewed' interactions	(4) & 'Settings reviewed' implies CRR value of 0
Choice of 'Generic'	Mean	0.311	0.311	0.311	0.311	0.311
	SD	0.463	0.463	0.463	0.463	0.463
Informed choice	Mean	0.500	0.500	0.500	0.500	0.500
	SD	0.500	0.500	0.500	0.500	0.500
Cost of 'Generic'	Mean	1.269	1.255	1.269	1.255	1.255
	SD	1.637	1.609	1.637	1.609	1.609
Cost x Informed	Mean	0.660	0.655	0.660	0.655	0.655
	SD	1.380	1.368	1.380	1.368	1.368
Q3 dummy	Mean	0.333	0.333	0.333	0.333	0.333
	SD	0.471	0.471	0.471	0.471	0.471
Q3 dummy x Informed	Mean	0.167	0.167	0.167	0.167	0.167
	SD	0.373	0.373	0.373	0.373	0.373
Not sure would make same choice in practice	Mean			0.144	0.140	0.140
	SD			0.351	0.347	0.347
Informed x Not sure	Mean			0.080	0.079	0.079
	SD			0.271	0.270	0.270
Cost x Not sure	Mean			0.279	0.268	0.268
	SD			0.925	0.908	0.908
Cost x Informed x Not sure	Mean			0.160	0.157	0.157
	SD			0.727	0.720	0.720
Q3 dummy x Not sure	Mean			0.072	0.070	0.070
	SD			0.258	0.255	0.255
Q3 dummy x Informed x Not sure	Mean			0.040	0.039	0.039
	SD			0.196	0.195	0.195
Settings reviewed	Mean				0.302	0.302
	SD				0.459	0.459
Informed x Reviewed	Mean				0.151	0.151
	SD				0.358	0.358
Cost x Reviewed	Mean				0.389	0.389
	SD				1.103	1.103
Cost x Informed x Reviewed	Mean				0.202	0.202
	SD				0.835	0.835
Q3 dummy x Reviewed	Mean				0.101	0.101
	SD				0.301	0.301
Q3 dummy x Informed x Reviewed	Mean				0.050	0.050
	SD				0.219	0.219

**Note:**

- Choice of 'Generic' (dependent variable) = 1 if 'Generic' (no data sharing) chosen; 0 otherwise
- Informed (choice) = 1 for Q1, Q2, Q3 of the set of informed choice; = 0 for Q1, Q2, Q3 of uninformed choices
- Cost of 'Generic' = £0 in Q1; = {£0.50, £1.00, £2.00, £3.00, £5.00} per month in Q2 and amount doubled or halved in Q3 depending on response to Q2; Cost of 'Generic' (no data sharing) relative to 'Targeted' (full data sharing)
- Not sure (would make same choice in practice) = 1 for Q2 and Q3 if the participant was 'Very sure' or 'Somewhat sure' that she/he would make the same choice in practice if the choice was really offered to her/him (follow-up question to Q2); = 0 if participant was 'Not very sure' or 'Not sure at all'
- (Settings) Reviewed = 1 if the participant had changed her/his data settings or had reviewed the settings but had chosen not to change them (unless this was because she/he wasn't sure how to); = 0 if the participant was aware that she/he could change the settings but had never gone in or wasn't aware or responded 'Don't know' or reviewed the settings but wasn't sure how to change them

**Table 26:** Sensitivity analysis and final model: Reward sample

Variable	(1) Baseline	(2) Baseline & 'Speeders' (fastest 10%) excluded	(3) Baseline & 'Not sure' interactions	(4) (1)&(2)&(3) & 'Settings reviewed' interactions	(5) (4) & 'Settings reviewed' implies CRR value of 0
Informed choice	1.091*** (0.072)	1.141*** (0.075)	0.999*** (0.074)	1.223*** (0.092)	1.223*** (0.092)
Cost of 'Generic'	-0.512*** (0.039)	-0.518*** (0.041)	-0.533*** (0.041)	-0.542*** (0.051)	-0.542*** (0.051)
Cost x Informed	0.205*** (0.035)	0.200*** (0.037)	0.210*** (0.038)	0.187*** (0.047)	0.187*** (0.047)
Q3 dummy	0.098 (0.071)	0.154** (0.074)	0.039 (0.078)	0.072 (0.096)	0.072 (0.096)
Q3 dummy x Informed	-0.046 (0.082)	-0.103 (0.085)	-0.032 (0.087)	-0.097 (0.106)	-0.097 (0.106)
Not sure would make same choice in practice			-2.477*** (0.240)	-2.564*** (0.252)	-2.564*** (0.252)
Informed x Not sure			1.674*** (0.295)	1.646*** (0.310)	1.646*** (0.310)
Cost x Not sure			0.427*** (0.079)	0.425*** (0.085)	0.425*** (0.085)
Cost x Informed x Not sure			-0.258*** (0.092)	-0.237** (0.098)	-0.237** (0.098)
Q3 dummy x Not sure			1.324*** (0.235)	1.441*** (0.244)	1.441*** (0.244)
Q3 dummy x Informed x Not sure			-0.863*** (0.307)	-0.922*** (0.323)	-0.922*** (0.323)
Settings reviewed				-0.006 (0.175)	-0.006 (0.175)
Informed x Reviewed				-0.586*** (0.163)	-0.586*** (0.163)
Cost x Reviewed				0.009 (0.081)	0.009 (0.081)
Cost x Informed x Reviewed				0.047 (0.078)	0.047 (0.078)
Q3 dummy x Reviewed				0.031 (0.165)	0.031 (0.165)
Q3 dummy x Informed x Reviewed				0.059 (0.197)	0.059 (0.197)
Constant	-0.111 (0.074)	-0.074 (0.078)	0.052 (0.078)	0.101 (0.097)	0.101 (0.097)
$\ln(\sigma_v^2)$	1.802*** (0.066)	1.812*** (0.069)	1.851*** (0.065)	1.864*** (0.068)	1.864*** (0.068)
Observations	12,084	11,070	12,084	11,070	11,070
Participants	2,014	1,845	2,014	1,845	1,845
Log likelihood	-6663	-6104	-6589	-6019	-6019
Pseudo R-squared	0.203	0.204	0.212	0.215	0.215
<b>Mean <math>\Delta E(\text{CS})</math> (£/user/month)</b>					
Uninformed	£2.26	£2.27	£2.36	£2.38	£1.68

**104 VALUE OF THE CHOICE REQUIREMENT REMEDY**

Uninformed: 95% CI	[£1.94-£2.58]	[£1.93-£2.61]	[£2.01-£2.71]	[£1.95-£2.81]	[£1.38-£1.99]
Informed	£5.77	£5.74	£5.66	£5.68	£4.03
Informed: 95% CI	[£4.87-£6.67]	[£4.83-£6.65]	[£4.72-£6.61]	[£4.61-£6.75]	[£3.26-£4.79]

Dependent variable: *choice*. Standard errors clustered by participant in parentheses. \* significant at 10%; \*\* significant at 5%; \*\*\* significant at 1%.  $\sigma_v$  denotes the standard deviation of the random effects. Mean change in expected consumer surplus calculated setting 'Not sure' = 0, 'Reviewed' = 0, 'Q3 dummy' = 0.

**Table 27:** Descriptive statistics – Reward sample

Variable		(1)	(2)	(3)	(4)	(5)
		Baseline	Baseline & 'Speeders' (fastest 10%) excluded	Baseline & 'Not sure' interactions	(1)&(2)&(3) & 'Settings reviewed' interactions	(4) & 'Settings reviewed' implies CRR value of 0
Choice of 'Generic'	Mean	0.479	0.487	0.479	0.487	0.487
	SD	0.500	0.500	0.500	0.500	0.500
Informed choice	Mean	0.500	0.500	0.500	0.500	0.500
	SD	0.500	0.500	0.500	0.500	0.500
Cost of 'Generic'	Mean	1.507	1.502	1.507	1.502	1.502
	SD	1.965	1.962	1.965	1.962	1.962
Cost x Informed	Mean	0.815	0.814	0.815	0.814	0.814
	SD	1.710	1.711	1.710	1.711	1.711
Q3 dummy	Mean	0.333	0.333	0.333	0.333	0.333
	SD	0.471	0.471	0.471	0.471	0.471
Q3 dummy x Informed	Mean	0.167	0.167	0.167	0.167	0.167
	SD	0.373	0.373	0.373	0.373	0.373
Not sure would make same choice in practice	Mean			0.109	0.110	0.110
	SD			0.312	0.313	0.313
Informed x Not sure	Mean			0.052	0.052	0.052
	SD			0.222	0.222	0.222
Cost x Not sure	Mean			0.247	0.249	0.249
	SD			1.001	1.004	1.004
Cost x Informed x Not sure	Mean			0.134	0.136	0.136
	SD			0.785	0.792	0.792
Q3 dummy x Not sure	Mean			0.054	0.055	0.055
	SD			0.227	0.228	0.228
Q3 dummy x Informed x Not sure	Mean			0.026	0.026	0.026
	SD			0.159	0.159	0.159
Settings reviewed	Mean				0.292	0.292
	SD				0.455	0.455
Informed x Reviewed	Mean				0.146	0.146
	SD				0.353	0.353
Cost x Reviewed	Mean				0.433	0.433
	SD				1.256	1.256
Cost x Informed x Reviewed	Mean				0.233	0.233
	SD				0.988	0.988
Q3 dummy x Reviewed	Mean				0.097	0.097
	SD				0.296	0.296
Q3 dummy x Informed x Reviewed	Mean				0.049	0.049
	SD				0.215	0.215

**Note:**

- Choice of 'Generic' (dependent variable) = 1 if 'Generic' (no data sharing) chosen; 0 otherwise
- Informed (choice) = 1 for Q1, Q2, Q3 of the set of informed choice; = 0 for Q1, Q2, Q3 of uninformed choices
- Cost of 'Generic' = £0 in Q1; = {£0.50, £1.00, £2.00, £3.00, £5.00} per month in Q2 and amount doubled or halved in Q3 depending on response to Q2; Cost of 'Generic' (no data sharing) relative to 'Targeted' (full data sharing)
- Not sure (would make same choice in practice) = 1 for Q2 and Q3 if the participant was 'Very sure' or 'Somewhat sure' that she/he would make the same choice in practice if the choice was really offered to her/him (follow-up question to Q2); = 0 if participant was 'Not very sure' or 'Not sure at all'
- (Settings) Reviewed = 1 if the participant had changed her/his data settings or had reviewed the settings but had chosen not to change them (unless this was because she/he wasn't sure how to); = 0 if the participant was aware that she/he could change the settings but had never gone in or wasn't aware or responded 'Don't know' or reviewed the settings but wasn't sure how to change them

## Appendix D – Literature Review Estimates

As discussed in Section 2.3, empirical studies related to the analysis and quantification of the value of online privacy include Hann et.al (2007), Savage and Waldman (2013) and Prince and Wallsten (2020). The estimates reported in these studies differ in various aspects such as the population studied, the dimensions or forms of online privacy valued as well as the unit of valuation reported by each of these studied. Hence the estimates reported by these studies are not directly comparable to the estimates reported by our study.

Hann et.al (2007) report values of U.S. subjects (upper-division undergraduate students from a major Eastern U.S. university) and Singaporean subjects (upper-division undergraduate students enrolled in an e-commerce technologies course at a major university) on various forms of privacy protection such as collection of information, error, unauthorized secondary use and improper access to online personal information on a per person per year basis.

Savage and Waldman (2013) report per person one-time willingness to pay values of US smartphone users to conceal their browser history, location data and to eliminate advertising.

Prince and Wallsten (2020) report willingness to accept values of users across different platforms (wireless carrier, financial institution, smartphone, and Facebook account) and across different countries (US, Mexico, Brazil, Colombia, Argentina and Germany) to share their bank balance information, their fingerprint information, to access and read their texts, to share information on their cash withdrawals and to share their location data on a per person per month basis.

In order to derive comparable estimates, we converted the values reported by the studies to British Pound Sterling (GBP) in 2021 prices (see Table 28). The values reported by Hann et.al (2007) were on a per person per year basis which we converted to a per person per month basis to make the units comparable.

A review of the values across the dimensions of privacy that are relevant to our study (i.e. collection and use/sharing of personal information online) as well as those that are comparable to the units of valuation of our study (i.e. per person per month basis) show that the value of online privacy reported in the literature vary from £1.44 to £2.46 per user per month. However, if we focus exclusively on the values reported for Facebook i.e. the willingness to accept payments from Facebook to allow it to share users' information with third parties, as reported by Prince and Wallsten (2020), we find that the values range from £2.70/month to £6.33/month to share access and read texts, £0.44/month to £5.98/month to share information about users' network and £0.41/month to £6.40/month to share contacts with Facebook. In general, the values for Facebook range from £0.41/month to £6.40/month.

Our preferred estimates of the value of the CRR, which range between range between £0.50 and £4.03 per user per month (see Section 5.3), are not out of line with the values reported in the literature on privacy valuation.

**Table 28:** Valuations of online privacy found in the literature

Reference	Form of online privacy valued	Sample	Unit of valuation	Values (in 2021 prices)
Hann et.al (2007)	collection, error, unauthorised secondary use & improper access to personal information online	US residents	per person per month	£1.68–£2.46
		Singapore residents	per person per month	£2.25
Savage & Waldman (2013)	to conceal browser history	US smartphone users	per person/one-time	£1.72
	to conceal location data		per person/one-time	£0.90
	to eliminate advertising		per person/one-time	£1.60
Prince & Wallsten (2020)	to share bank balance	Users across platforms (wireless carriers, financial institution, smartphone and Facebook) and countries (Argentina, Brazil, Colombia, Germany, Mexico and US)	per person per month	£6.67
	to share fingerprint information		per person per month	£5.98
	to access and read texts		per person per month	£4.78
	to share information on cash withdrawals		per person per month	£4.59
	to share location data		per person per month	£1.44
	read texts	Argentinian Facebook users	per person per month	£5.42
	share information about network		per person per month	£1.61
	share contacts		per person per month	£0.92
	read texts	Brazilian Facebook users	per person per month	£2.70
	share information about network		per person per month	£0.44
	share contacts		per person per month	£0.51
	read texts	Colombian Facebook users	per person per month	£5.23
	share information about network		per person per month	£0.90
	share contacts		per person per month	£0.41
	read texts	German Facebook users	per person per month	£6.33
share information about network	per person per month		£5.98	
share contacts	per person per month		£6.40	
read texts	Mexican Facebook users	per person per month	£5.14	
share information about network		per person per month	£1.25	
share contacts		per person per month	£1.82	
read texts	US Facebook users	per person per month	£3.88	
share information about network		per person per month	£2.27	
share contacts		per person per month	£2.81	

# Appendix E – Review by Prof. Ken Willis

## Value of the Choice Requirement Remedy.

Final Report to Which?

June 2021

by

**Accent and PJM Economics**

Reviewed by Ken Willis

Newcastle University

10<sup>th</sup> July 2021

### Objective

Google and Facebook enjoy substantial market power in the search and display advertising. It is argued that this results in consumers not receiving adequate compensation for the use of their personal data by online platforms and for the attention consumers devote to advertising material.

The study investigates the Competition and Markets Authority (CMA) recommendation that the Government introduces legislation requiring platforms to give consumers the choice of not to share their data for personalised advertising. This is the Choice Requirement Remedy (CRR). Consumers who choose not to share their personal data would continue to see generic contextual adverts in the service from the online platform. But consumers could be provided with incentives to accept targeted personal advertising, which might benefit consumers and help online platforms manage their revenue. The objective of the study is to value this CRR.

The study provides a very thorough and commendable literature review on the value that consumers assign to privacy and personal data, as a prelude to devising a framework and methodology to value CRR.

### Methodology

PJM Economics and Accent rightly adopt a stated preference valuation framework based on offering respondents a choice between two alternatives (1) full sharing of personal data with targeted adverts; and (2) no sharing on personal data with generic adverts. The value of these options was estimated through (a) the maximum that consumers would be willing-to-pay (WTP) (a fee) for no sharing of personal data, and (b) minimum consumers would be willing-to-accept (WTA) (a reward e.g. PayPal credit or Amazon vouchers) for full sharing of personal data and receiving targeted adverts.

The alternative method would have been to use a revealed preference, based on actual users' decisions on the use of their information. But this was not feasible since the vast majority of

Google and Facebook users do not, for various reasons, amend their privacy settings to prevent personalized advertising, despite having the option to do so. Moreover, where users have reviewed their settings, this decision provided no information on the value of CRR.

Since the choice is a dichotomous one between the use of personalised information to generate personalised adverts versus no personalised information use and generic adverts being presented, the study appropriately adopts a discrete choice (DC) contingent valuation (CV) method. A CV method, presents respondents with two alternative scenarios, from which respondents choose their most preferred. CV can adopt different payment forms (open ended WTP question, payment card, iterative bidding, dichotomous choice, etc). But a single bounded (SB) DC CV has been shown to be the most incentive compatible; whilst a double bounded (DB) DC CV is statistically more efficient, and usually produces more conservative estimates of WTP than a SB DC CV. However, a DB DC CV may, as the Report states, be subject to anchoring bias based on the first bid amount, although bound or path inconsistencies have been shown to be insignificant where bid amounts are small (as in the present study) compared to large WTP bid amounts.

The study also was innovative in adapting CV to derive values of different types of privacy protection, through presenting respondents with 5 choice scenarios each with different personal information settings and price amounts.

### Survey

The survey follows good practice in the application of stated preference questionnaire, with screening questions to ensure a representative sample, respondent information on advertising awareness, followed by choice questions, questions to assess the validity of responses, and ending with respondent's socio-demographic information. Following good practice, the questionnaire was also thoroughly tested through qualitative in-depth interviews with respondents, followed by a large pilot survey.

The main survey of 4,014 interviews is large compared to most surveys in studies using SB and DB DC CV methods, and thus would *a priori* be expected to produce statistically accurate, reliable, and robust results. The Report appropriately includes a caveat pointing out the sample comprises respondents who have signed up to complete surveys on a regular basis, and thus may be more likely to share data about themselves compared to respondents in the general population; in which case the valuation estimates may underestimate the true population valuation for CRR.

The study is commendable in testing the validity of responses through follow-up questions, testing for ordering effects in question presentation, and testing the sensitivity of results by excluding those respondents who completed the questionnaire excessively quickly.

### Econometric estimation

The Report proves an excellent description of the approach to the econometric modelling of the CV data, in Appendix C: Econometric Modelling of Contingent Valuation Choices and CRR Valuation.

The random effect models fit the data quite well as exemplified by the pseudo- $R^2$  statistics; and nearly all the variables are statistically significant. The models produce plausible CRR valuation results, with tight confidence limits.

Parametric logit models and random effects logit models are used to estimate CV CRR values. These models rely on assumptions about the underlying distribution of the data. The data also show, in Figures 25 to 27, well behaved downward sloping demand curves, which support the parametric CV model results reported.

The validity of CV studies can be judged in terms of content validity (has the good being valued been described clearly; is the payment vehicle realistic); convergent validity (are the CV estimates similar to those produced by other techniques); theoretical validity (are results consistent with theoretical expectations); and criterion validity (are behaviour and values in experimental markets consistent with behaviour and prices in actual markets).

Content validity seems to have been satisfied: respondents understood the issue and were confident in their CV choices. The study appropriately compares the CRR values with values for comparable types of privacy and information goods. Although it notoriously problematic to compare stated and revealed preference values across countries and cultures, similar ballpark estimates were noted. So convergent validity might be deemed to be satisfied. The results are also consistent in terms of theoretical expectations, e.g. the coefficients for the cost variables are negative; WTP values for no sharing of information were less than WTA compensation values for sharing of personal information. Criterion validity is perhaps the ultimate yardstick against which to judge the validity of responses, but rarely is it possible to judge the truthfulness of SP estimates on this basis.

### **Conclusion**

The “*Value of the Choice Requirement Remedy*” study has been expertly undertaken by Accent and PJM Economics. The methodology, questionnaire, and the stated preference contingent valuation questions, worked well. Respondents could clearly understand the tasks required, and gave valid responses to questions asked. The econometric analysis of the data was skilfully and expertly implemented and produced statistically significant estimates of the Choice Requirement Remedy (CRR). The WTP values for CRR appear to be accurate, and seem to be intuitively reasonable. The impact of attributes and WTP values on demand for conform to economic theory.

The “*Value of the Choice Requirement Remedy*” study by Accent and PJM Economics is a meticulous, proficient, and professional piece of research. Which? can have confidence in the results.

Which?, 2 Marylebone Road,  
London NW1 4DF  
Phone +44 (0)20 7770 7000  
Fax +44 (0)20 7770 7600